Assiut University

Faculty of Science

Department of Mathematics

Course Title: Discrete Structures

Course Code: CS201

Course hours per week:

Lecture	Tutorial / Practical	Total
3	2	5

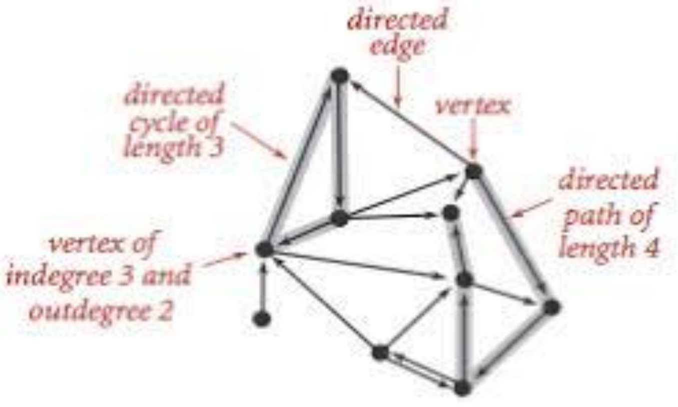
For

Faculty of Computers & Information

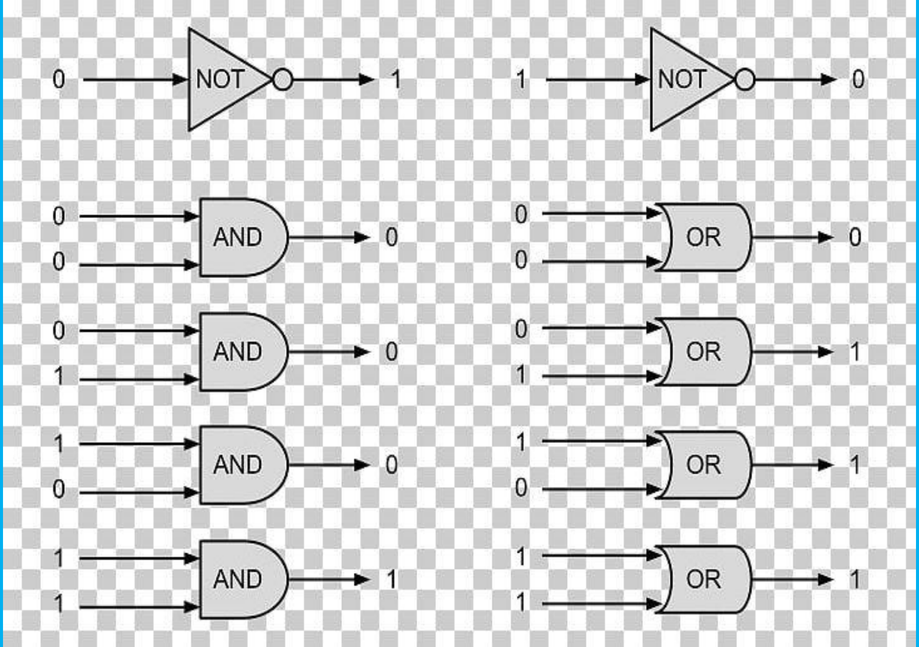
By

Professor

OSAMA RASHED SAYED


directed edge
directed cycle of length 3
vertex
directed path of length 4
vertex of indegree 3 and outdegree 2

Anatomy of a digraph



Contents

Contents

CHAPTER (I): SETS, RELATIONS AND FUNCTIONS 7

1.1 Sets 7

Subsets, Combining sets: union and intersection, Venn diagrams, Universal set, The power set, Disjoint sets, Cartesian product, Cardinality, Comparable sets, Multiset, Multiplicity, Set operations, The complement, Set identities, Computer representation of the sets, Exercise set (1.1).

1.2 Functions 38

What is a Function?, Set of ordered pairs, Piecewise functions, The absolute value function, Floor and ceiling Functions, Domain, Codomain and Range, Injective, Surjective and Bijective, Inverse function, Operations with functions, Composition of functions, De-Composing function, Cardinality, Countable set, Exercise Set (1.2).

1.3 Relations 89

Functions as relations, Properties of relations, Operations on binary relations, Equivalence relations, Equivalence classes, Representing relations using matrices, Representing relations using digraphs, Exercises Set (1.3).

CHAPTER (II): MATHEMATICAL LOGIC 135

2.1 Propositional Calculus 135

Logical Connectives, Order of precedence, Tautologies and contradictions, Logical Equivalence, Exercise Set (2.1).

2.2 Predicates and Quantifiers 158

Predicates, Quantifiers, Logical Equivalence Involving Quantifiers, Nested Quantifiers, The Order of Quantifiers, Solved problems, Exercise Set (2.2).

2.3 Rules of inference 226

Arguments and Validity, Methods to Test Validity of an Argument, Rules of Inference for Propositional Logic, Using Rules of Inference to Build Arguments, Resolution and Automated Theorem Proving, Fallacies, Rules of Inferences for Quantified Statements, Combining Rules of Inference for Propositions and Quantified Statements, Exercise Set (2.3).

CHAPTER (III): METHODS OF PROOF 276

3.1 Introduction 276

Some Terminology.

3.2 Methods of Proving Theorems 277

Direct proofs, Proof by contraposition (indirect proof), Proof by contradiction, Proofs of equivalent, Disproving universal statements, Proof by cases, Existence proofs, Exercises Set (3).

CHAPTER (IV): MATHEMATICAL INDUCTION 299
4.1 The Basic Principle 299
4.2 Variations of the Basic principle 302
4.3 Miscellaneous Examples 311
4.4 Higher Dimensional Induction 316
Exercise Set (4) 320.
CHAPTER (V): ELEMENTARY NUMBER THEORY 323
5.1 The Ring of Integers 323
5.2 Prime Numbers 332
5.3 Divisibility 345
5.4 The Division Algorithm 349
5.5 Greatest Common Divisors 350
5.6 The Euclidean Algorithm 356
5.7 gcds as Linear Combinations 363
5.8 The Fundamental Theorem of Arithmetic 367
5.9 The Chinese Remainder Theorem 374
Exercise Set (5) 387
CHAPTER (VI): COUNTING 391
6.1 The Basics of Counting 391
6.2 Basic Counting Principles 392
1-The Addition rule or Sum rule, 2-Product Rule (The Multiplication Rule), 3-The Subtraction Rule (Inclusion–Exclusion for Two Sets), 4-Tree diagrams, 5-Pigeonhole Principle, 6-generalized Pigeonhole Principle, 7-Permutations and Combinations, Combinatorial Proofs.
Exercise (6) 448
CHAPTER (VII): BOOLEAN ALGEBRA AND THEIR APPLICATIONS 456
7.1 Boolean Algebra 456
7.2 Boolean Functions 468
7.3 Logic Gates 479
Exercises set (7) 491

CHAPTER (VIII): GRAPH THEORY 494

8.1 Introduction 494

8.2 Graphs and Graph Models 495

8.3 Subgraph 514

8.4 Special Graphs 517

8.5 Representation of Graphs 523

8.6 Isomorphism of Graphs 532

8.7 Connected Graphs 541

8.8 Planar graph 549

Euler's formula, Graph coloring, Euler paths,

Exercises Set (8) 568

CHAPTER (IX): TREES 573

9.1 Tree 573

9.2 Examples of Trees 578

Forest, A rooted tree, Spanning Tree, Star Graph, Banana Tree, Centered Tree, Bi-centered Tree, Binary Tree, Red-Black Tree.

Exercise Set (9) 591

Bibliography 594

CHAPTER (I)

SETS, RELATION AND FUNCTIONS

Chapter (I)

Sets, Relations and Functions

Much of discrete mathematics is devoted to the study of discrete structures, used to represent discrete objects. Many important discrete structures are built using sets, which are collections of objects. Relations between elements of sets occur in many contexts. Every day we deal with relationships such as those between a business and its telephone number, a person and a relative and so on. In mathematics we study relationships such as these between a positive integer and one that it divides, an integer and one that it is congruent to modulo 5, and so on. The concept of a function is extremely important in discrete mathematics. A function assigns to each element of a set exactly one element of a set. Functions play important roles throughout discrete mathematics.

1.1 Sets

Definition.

A ***set*** is an unordered collection of objects, called *elements* or *members* of the set. A set is said to *contain* its elements. We write $a \in A$ to denote that a is an element of the set A. The notation $a \in A$ denotes that a is not an element of the set A.

It is common for sets to be denoted using uppercase letters. Lowercase letters are usually used to denote elements of sets. **There are several ways to describe a set**. One way is to list all the members of a set, when this is possible. We use a notation where all members of the set are listed between braces. For example, the notation $\{a, b, c, d\}$ represents the set with the four elements *a*, *b*, *c*, and *d*. This way of describing a set is known as the **roster method**.

Example

The set *V* of all vowels in the English alphabet can be written as $V = \{a, e, i, o, u\}$. ■

Example

The set *O* of odd positive integers less than 10 can be expressed by $O = \{1, 3, 5, 7, 9\}$. ■

Although sets are usually used to group together elements with common properties, there is nothing that prevents a set from having seemingly unrelated elements.

Example

{*a*, 2, Ali, Assiut} is the set containing the four elements *a*, 2, Ali, and Assiut. ■

If we can completely list (enumerate) all the elements in a set, the set is said to be **finite**. The set of primary colours is finite set. If a set isn't finite, it is said to be **infinite**. The set of all positive integers is an infinite set.

Sometimes the roster method is used to describe a set without listing all its members. Some members of the set are listed, and then *ellipses* (. . .) are used when the general pattern of the elements is obvious.

Example

The set of positive integers less than 100 can be denoted by

$\{1, 2, 3, \ldots, 99\}$. ■

Another way to describe a set is to use **set builder** notation. We characterize all these elements in the set by stating the property or properties they must have to be members.

Example

The set O of all odd positive integers less than 10 can be written as

$O = \{x \mid x \text{ is an odd positive integer less than } 10\}$.

or, specifying the universe as the set of positive integers, as

$O = \{x \in \mathbb{Z}^+ \mid x \text{ is odd and } x < 10\}$. ■

We often use this type of notation to describe sets when it is impossible to list all the elements of the set.

Example

The set $\mathbb{Q}^+$ of all positive rational numbers can be written as $\mathbb{Q}^+ = \{x \in R \mid x = p/q, \text{for some positive integers } p \text{ and } q\}$.

Here are some common mathematical sets you are familiar with. You need to be able to recognize the symbols.

$\mathbb{N}$ the set of positive integers and zero,

$\mathbb{Z}$ the set of integers

$\mathbb{Z}^+$ the set of positive integers

$\mathbb{Q}$ the set of rational numbers

$\mathbb{Q}^+$ the set of positive rational numbers

$\mathbb{R}$ the set of real numbers

$\mathbb{R}^+$ the set of positive real numbers

$\mathbb{C}$ the set of complex numbers

(Note that some people do not consider 0 a natural number, so be careful to check how the term *natural numbers* is used when you read different books.)

Recall the notation for **intervals** of real numbers. When *a* and *b* are real numbers with $a < b$, we write

$$[a, b] = \{x \mid a \leq x \leq b\}$$
$$[a, b) = \{x \mid a \leq x < b\}$$
$$(a, b] = \{x \mid a < x \leq b\}$$
$$(a, b) = \{x \mid a < x < b\}$$

Note that $[a, b]$ is called the **closed interval** from *a* to *b* and (a, b) is called the **open interval** from *a* to *b*.

Sets can have other sets as members, as the following example illustrates.

Example.

The set $\{\mathbb{N}, \mathbb{Z}, \mathbb{Q}, \mathbb{R}\}$ is a set containing four elements, each of which is a set. The four elements of this set are $\mathbb{N}$, the set of natural numbers; $\mathbb{Z}$, the set of integers; $\mathbb{Q}$, the set of rational numbers; and $\mathbb{R}$, the set of real numbers. ■

Remark

Note that the concept of a datatype, or type, in computer science is built upon the concept of a set. In particular, a **datatype** or **type** is the name of a set, together with a set of operations that can be performed on objects from that set.

Example

Boolean is the name of the set $\{0, 1\}$ together with operators on one or more elements of this set, such as AND, OR, and NOT.■

● Definition of equality for sets

Two sets S and T are equal if every element of S is also an element of T and every element of T is also an element of S. Not surprisingly, write as S = T.

Here are some implications of this definition.

● The ordering of elements in a set is not important

The set {red, yellow, blue} equals (i.e. is the same as) the set {yellow, blue, red}. Why? Look at the definition of equality. Every element in the first set is an element of the second, and every element in the second set is an element of the first. So the two sets are equal.

● Something is either an element of a set or not; it doesn't make any difference if you list it multiple times

The set {red, red, yellow, blue, red} is the same as (i.e. is equal to) the set {red, yellow, blue}, even though “red” is listed multiple times in the first set. Don’t take my word for it; check the definition of equals.

●The smallest possible set

We call the set containing no elements the **null set** or the **empty set**. It sometimes is written as { } but more often we write it as ϕ. For instance, the set of all positive integers that are greater than their squares is the null set.

A set with one element is called a **singleton set**. A common error is to confuse the empty set ϕ with the set $\{\phi\}$, which is a singleton set. The single element of the set $\{\phi\}$ is the empty set itself! A useful analogy for remembering this difference is to think of folders in a computer file system. The empty set can be thought of as an empty folder and the set consisting of just the empty set can be thought of as a folder with exactly one folder inside, namely, the empty folder.

● Subsets

Definition

The set A is a *subset* of B if and only if every element of A is also an element of B. We use the notation $A \subseteq B$ to indicate that A is a subset of the set B.

We see that $A \subseteq B$ if and only if the quantification

$$\forall x(x \in A \rightarrow x \in B)$$

is true.

We say that S is **a superset** of T if every element of T is an element of S. We write this as S ⊇ T.

To show that $A \subseteq B$, show that if x belongs to A then x also belongs to B.

To show that $A \nsubseteq B$, find a single $x \in A$ such that $x \notin B$.

Example.

The set of all odd positive integers less than 10 is a subset of the set of all positive integers less than 10, the set of rational numbers is a subset of the set of real numbers, the set of all computer science majors at your school is a subset of the set of all students at your school, and the set of all people in Egypt is a subset of the set of all people in Egypt (that is, it is a subset of itself). Each of these facts follows immediately by noting that an element that belongs to the first set in each pair of sets also belongs to the second set in that pair. ■

Example.

(i) The set of integers with squares less than 100 is not a subset of the set of nonnegative integers because -1 is in the former set [as $(-1)^2 < 100$], but not the later set.

(ii) The null set is a subset of every set, i.e., If A is any set then $\phi \subset A$. ■

●The relationship between subsets, supersets and equality

Reviewing the definitions, we see that for two sets S and T, S = T is true whenever both S ⊆ T and S ⊇ T are true.

To show that two sets A and B are equal, show that $A \subseteq B$ and $B \subseteq A$.

Sets may have other sets as members. For instance, the sets $A = \{\emptyset, \{a\}, \{b\}, \{a, b\}\}$ and $B = \{x \mid x$ is a subset of the set $\{a, b\}\}$. Note that these two sets are equal, that is, $A = B$. Also note that $\{a\} \in A$, but $a \notin A$.

● "Proper" subsets

Sometimes we have S ⊆ T and we want to rule out the possibility that S = T. To do this, we write S ⊂ T, i.e. we omit the bar below the ⊂. To say this in words, we say that S is a proper subset of T. The use of the word "proper" here is kind of funny. It is just the term that mathematicians have come to use to avoid having to say, "S is a subset of T but it isn't equal to T." Similarly, S ⊃ T is read “S is a proper superset of T ” and is a shorter way of writing S ⊇ T and S ≠ T.

Example

(i) If $A = \{0, 2, 9\}$, $B = \{0, 2, 7, 9, 11\}$, then $A \subset B$ (A is a proper subset of B).

(ii) If $A = \{a, a, b\}$, $B = \{a, b\}$, then A and B denoted the same set, i.e., $A = B$.

(iii) If $A = \{1, 2, 4\}$, $B = \{2, 4, 6, 8\}$, then A is proper subset of B and B is a superset of A.■

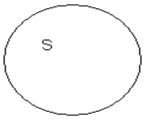
●Combining sets: union and intersection

So far, we have defined various relations on pairs of sets (=, ⊆, ⊂ etc.) in terms of membership. It is also useful to define operations that take two sets and form a third set. Once again, we will define these operations in terms of membership. We’ll start by defining the **intersection** of two sets S and T to be the set containing anything that is *both* an element of S *and* an element of T. We’ll write the intersection operation as S ∩ T.

Two sets are called disjoint if their intersection is the empty set. Similarly, we’ll define the **union** of two sets S and T to be the set containing anything that is *either* an element of S *or* an element of T. We’ll write the union operation as S ∪ T.

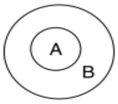
●Venn diagrams

For elementary set operations, there is a conventional method of drawing pictures called **Venn diagrams**, named after the British mathematician John Venn. To draw a set S, we simply draw a circle, with the name of the set inside the circle.

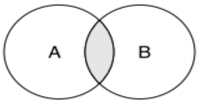


The intent of this drawing is that the inside of the circle represents all the elements in S. The outside of the circle represents everything that isn’t in the set S. There isn’t any significance to the fact that we use circles in Venn diagrams. We could just as well draw

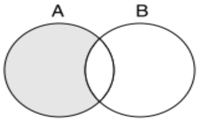
Now, to represent an operation on two sets, we draw two overlapping circles, like this:



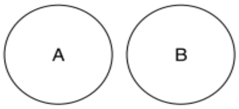
$A \subset B$ (A is a subset of B)



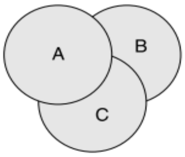
$A \cap B$ (Shaded part)



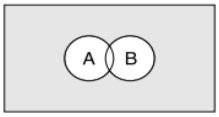
$A - B$ (Shaded part)



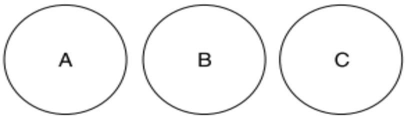
$A \cap B = \phi$



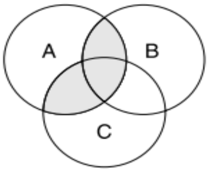
$A \cap (B \cup C)$



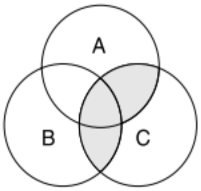
$(\overline{A \cup B})$ (Shaded)



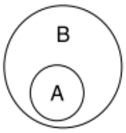
$A \cap (B \cap C) = \phi$



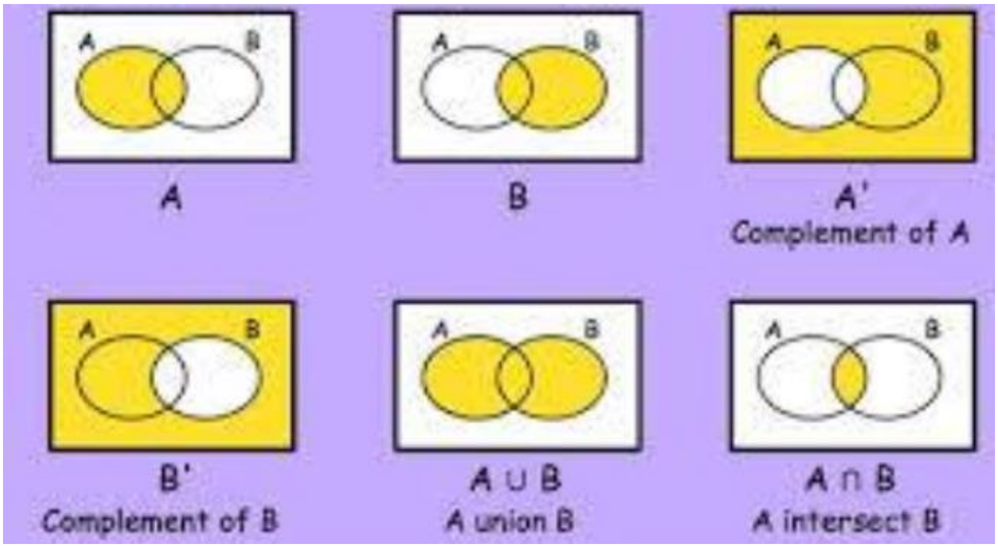
$(A \cap B) \cup (A \cap C)$ (Shaded part)



$(A \cup B) \cap C$ (Shaded)



$A \cap B = A$ if $A \subseteq B$



Draw a Venn diagram (not limited to circles) that depicts every possible combination of intersections between four sets. What is the best you can do?

●Universal Set

Definition.

In many discussions all the sets are considered to be subsets of one particular set. This set is called the universal set for that discussion. The Universal set is often designated by the script letter *U* (or by *X*). Universal set in not unique, and it may change from one ***discussion*** to another.

Example.

If $A = \{0, 2, 7\}$, $B = \{3, 5, 6\}$, $C = \{1, 8, 9, 10\}$, the universal set can be taken as the set. $U = \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10\}$.■

●The Power Set

The **power set** of S is the set of all subsets of the set S. The power set of S is denoted by $P(S)$.

Example.

What is the power set of the set {0,1,2}?

Solution.

$P(\{0,1,2\}) = \{\phi, \{0\}, \{1\}, \{2\}, \{1,2\}, \{0,2\}, \{0,1\}, \{0,1,2\}\}$.■

If a set has n elements, then its power set has 2^n elements.

Example.

What is the power set of the empty set? What is the power set of the set $\{\phi\}$?

Solution.

The empty set has exactly one subset, namely, itself.

Consequently,

$$P(\phi) = \{\phi\}.$$

The set $\{\phi\}$ has exactly two subsets, namely, ϕ and the set $\{\phi\}$ itself. Therefore, $P(\{\phi\}) = \{\phi, \{\phi\}\}$. ■

● Disjoint Sets

Definition.

Two sets are said to be disjoint if they have no element in common.

Example.

The sets $A = \{0, 4, 7, 9\}$ and $B = \{3, 6, 10\}$ are disjoint. ■

●Cartesian Product

The order of elements in a collection is often important. Because sets are unordered, a different structure is needed to represent ordered collections. This is provided by **ordered *n*-tuples**.

Definition.

The *ordered n-tuple* $(a_1, a_2, \ldots, a_n)$ is the ordered collection that has a_1 as its first element, a_2 as its second element, . . . , and a_n as its *n*th element. We say that two ordered *n*-tuples are equal if and only if each corresponding pair of their elements is equal. In other words, $(a_1, a_2, \ldots, a_n) = (b_1, b_2, \ldots, b_n)$ if and only if $a_i = b_i$ for $i = 1, 2, \ldots, n$. In particular, ordered 2-tuples are called **ordered pairs**. The ordered pairs (a, b) and (c, d) are equal if and only if $a = c$ and $b = d$. Note that (a, b) and (b, a) are not equal unless $a = b$.

Definition.

Let A and B be sets. The **Cartesian product** of A and B, denoted by $A \times B$, is the set of all ordered pairs (a, b) where $a \in A$ and $b \in B$. Hence

$$A \times B = \{(a, b): a \in A \land b \in B\}$$

Example.

What are the Cartesian products $A \times B$ and $B \times A$, where $A = \{1, 2\}$ and $B = \{a, b, c\}$?

Solution.

$A \times B = \{(1, a), (1, b), (1, c), (2, a), (2, b), (2, c)\}$.

$B \times A = \{(a, 1), (a, 2), (b, 1), (b, 2), (c, 1), (c, 2)\}$. ■

Note that the Cartesian product $B \times A$ is not equal to the Cartesian product $A \times B$.

Definition.

The Cartesian product of the sets $A_1, A_2, \dots, A_n$, denoted by $A_1 \times A_2 \times \dots \times A_n$, is the set of ordered n-tuples $(a_1, a_2, \dots, a_n)$, where $a_i \in A_i$ for $i = 1,2, \dots, n$. In other words,

$$A_1 \times A_2 \times \dots \times A_n = \{(a_1, a_2, \dots, a_n): a_i \in A_i, i = 1,2, \dots, n\}$$

Example.

What is the Cartesian product $A \times B \times C$, where $A = \{0, 1\}$, $B = \{1, 2\}$, and $C = \{0, 1, 2\}$?

Solution.

The Cartesian product $A \times B \times C$ consists of all ordered triples (a, b, c), where $a \in A$, $b \in B$, and $c \in C$.

Hence,

$$A \times B \times C = \{(0,1,0),(0,1,1),(0,1,2),(0,2,0),(0,2,1),(0,2,2),(1,1,0),(1,1,1),(1,1,2),(1,2,0),(1,2,1),(1,2,2)\}.$$ ■

Remark.

Note that when A, B, and C are sets, $(A \times B) \times C$ is not the same as $A \times B \times C$.

Definition.

A subset R of the Cartesian product $A \times B$ is called a **relation** from the set A to the set B. The elements of R are ordered pairs, where the first element belongs to A and the second to B. For example $R = \{(a,0),(a,1),(a,3),(b,1),(b,2),(c,0),(c,3)\}$ is a relation from the set $\{a,b,c\}$ to the set {0, 1, 2, 3}.

A relation from a set A to itself is called a relation on A.

Example.

What are the ordered pairs in the less than or equal to relation, which contains (a,b) if $a \leq b$, on the set $\{0,1,2,3\}$?

Solution.

The ordered pair (a,b) belongs to R if and only if both a and b belong to $\{0,1,2,3\}$ and $a \leq b$. Consequently, the ordered pairs in R are $(0,0), (0,1), (0,2), (0,3), (1,1), (1,2), (1,3), (2,2), (2,3)$, and $(3,3)$. ■

We will study relations and their properties in Section 1.3.

●Cardinality

Sets are used extensively in counting problems, and for such applications we need to discuss the sizes of sets.

●Cardinality of a Set

Definition.

Let S be a set. If there are exactly n distinct elements in S where n is a nonnegative integer, we say that S is a *finite set* and that n is the *cardinality* of S. The cardinality of S is denoted by $|S|$.

Example.

Let A be the set of odd positive integers less than 10.

Then $|A| = 5$.■

Example.

Let S be the set of letters in the English alphabet.

Then $|S| = 26$. ■

Example.

Because the null set has no elements, it follows that $|\phi| = 0$. ■

We will also be interested in sets that are not finite.

Definition.

A set is said to be *infinite* if it is not finite.

Example.

The set of positive integers is infinite. ■

●Cardinality of Union of Two Sets

To find the number of elements in the union of two finite sets *A* and *B*, not that $|A| + |B|$ counts each element that is in *A* but not in *B* or in *B* but not in *A* exactly once, and each element that is in both *A* and *B* exactly twice. Thus, if the number of elements that are in both *A* and *B* is subtracted from $|A| + |B|$, elements in $A \cap B$ will be counted only once.

Hence $|A \cup B| = |A| + |B| - |A \cap B|$.

●Cardinality of Union of Three Sets

Number of elements in $A \cup B \cup C$: If *A*, *B* and *C* are any three finite sets, then

$|A \cup B \cup C| = |A| + |B| + |C| - |A \cap B| - |A \cap C| - |B \cap C| + |A \cap B \cap C|$.

● Comparable Sets

Definition.

Two sets *A* and *B* are said to be comparable if $A \subset B$ or $B \subset A$.

Definition.

Two sets *A* and *B* are said to be comparable if $A \not\subset B$ and $B \not\subset A$.

Example.

Let $A = \{1, 2, 3\}$ and $B = \{1, 2, 3, 4, 6\}$ then A is comparable to B, since A is a subset of B.■

Example.

If $A = \{a, c\}$, $B = \{b, c, d, e, f\}$ then $A \not\subset B$ and $B \not\subset A$. Therefore the sets A and B are not comparable. ■

●Multiset

Definition.

A collection of objects that are not necessarily distinct is called a multiset.

Example.

$\{a, a, b, b\ c, c\}$.■

●Multiplicity

Definition.

Let S be a multiset and $x \in S$. The multiplicity of x is defined to be the numbers of times the element x appears in the multiset S.

Example

Let $S = \{a, a, b, b, b, d, d, d, e\}$. Then

Multiplicity of a is 2;

Multiplicity of b is 3;

Multiplicity of d is 3;

Multiplicity of e is 1.

If A and B are multisets then $A \cup B$ and $A \cap B$ are also multisets. The multiplicity of an element $x \in A \cup B$ is equal to the maximum of the multiplicity of x in A and B. The multiplicity of $x \in A \cap B$ is equal to the minimum of the multiplicities of x in A and in B.

Example.

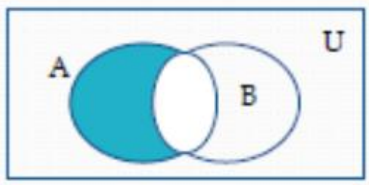
Let $A = \{a, a, a, b, b, c, c, d, d\}$ and $B = \{a, a, b, c, d\}$. Then $A \cup B = \{a, a, a, b, b, c, c, d, d\}$ and $A \cap B = \{a, a, b, c, d\}$. ■

●Set Operations

Let A and B be sets. The **difference** of A and B, denoted by $A – B$, is the set containing those elements that are in A but not in B.

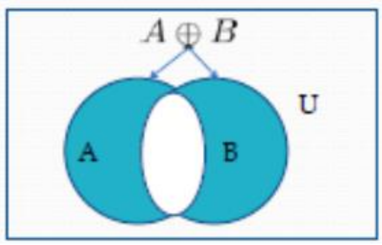
The difference of A and B is also called the **complement of *B* with respect to *A***.

Thus $A - B = \{x : x \in A \wedge x \notin B\} = A \cap B^c$.



Venn Diagram for A – B

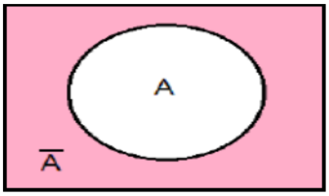
The **symmetric difference** of A and B, denoted by $A \oplus B$, is defined as $A \oplus B = A \cup B - A \cap B = (A - B) \cup (B - A)$.



For example, $\{1,3,5\} - \{1,2,3\} = \{5\}$, $\{1,2,3\} - \{1,3,5\} = \{2\}$ and $\{1,35\} \oplus \{1,2,3\} = \{5\} \cup \{2\} = \{2,5\}$.

● The complement

Let U be the universal set. The **complement** of the set A, denoted by A^c (or $\bar{A}$), is the complement of A with respect to U. In other words, the complement of the set A is $U - A$. An element belongs to A^c if and only if $x \notin A$. This tells us that $A^c = \{x: x \notin A\}$.



Once the universal set U has been specified, the complement of a set can be defined.

●Set Identities

The following table lists the most important set identities. We will prove several of these identities here, using three different methods.

Identity	Name
$A \cup \phi = A$ $A \cap U = A$	Identity laws
$A \cup U = U$ $A \cap \phi = \phi$	Domination laws
$A \cup A = A$ $A \cap A = A$	Idempotent laws
$(A^c)^c = A$	Complementation law
$A \cup B = B \cup A$ $A \cap B = B \cap A$	Commutative laws
$A \cup (B \cup C) = (A \cup B) \cup C$ $A \cap (B \cap C) = (A \cap B) \cap C$	Associative laws
$A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$ $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$	Distributive laws
$(A \cup B)^c = A^c \cap B^c$ $(A \cap B)^c = A^c \cup B^c$	De Morgan’s laws
$A \cup (A \cap B) = A$ $A \cap (A \cup B) = A$	Absorption laws
$A \cup A^c = U$ $A \cap A^c = \phi$	Complement laws

Example.

We will prove that $(A \cap B)^c = A^c \cup B^c$ by showing that each is a subset of the other. First suppose that $x \in (A \cap B)^c$. By the definition of complement, $x \notin A \cap B$. Hence, $x \notin A$ or $x \notin B$. By the definition of the complement $x \in A^c$ or $x \in B^c$. By the definition of the union that $x \in A^c \cup B^c$. So $(A \cap B)^c \subseteq A^c \cup B^c$. Now, suppose that $x \in A^c \cup B^c$. By the definition of union, $x \in A^c$ or $x \in B^c$. Hence, $x \notin A$ or $x \notin B$. By the definition of complement, $x \notin A \cap B$. It follows $x \in (A \cap B)^c$. This shows that $A^c \cup B^c \subseteq (A \cap B)^c$. Since we have shown that each set is a subset of the other, the two sets are equal, and the identity is proved.■

Example.

We will use **set builder notation** and logical equivalence to show that $(A \cap B)^c = A^c \cup B^c$ as follows:

$$(A \cap B)^c = \{x: x \notin A \cap B\} = \left\{x: \neg\left(x \in (A \cap B)\right)\right\}$$

$$= \{x: \neg(x \in A \wedge x \in B)\} = \{x: x \notin A \vee x \notin B\}$$

$$= \{x: x \in A^c \vee x \in B^c\} = \{x: x \in A^c \cup B^c\}$$

$$= A^c \cup B^c. \blacktriangleleft$$

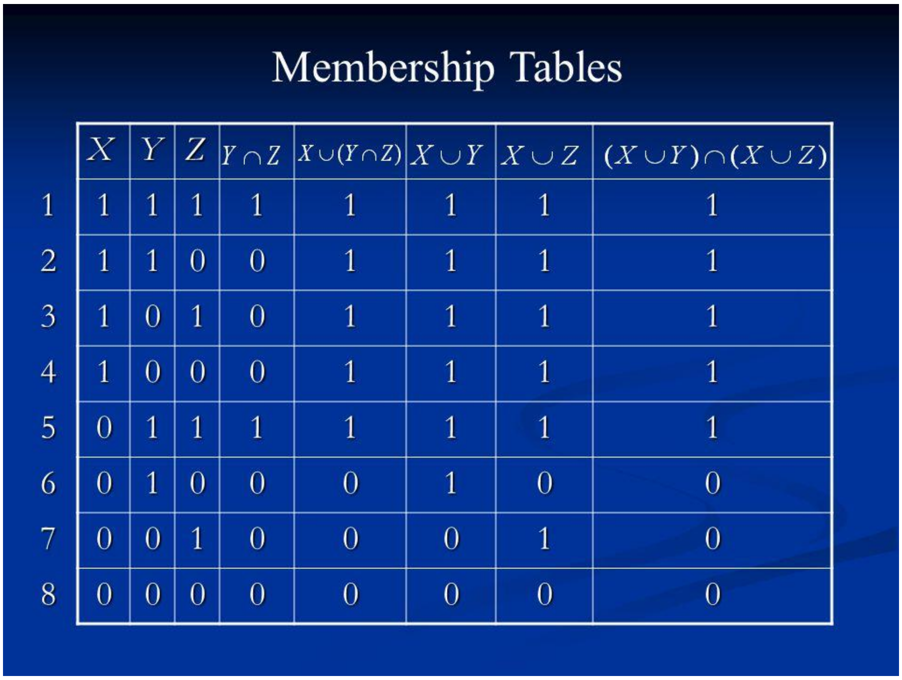
Set identities can also be proved using **membership tables.** We consider each combination of sets that an element can belong to and verify that elements in the same combinations of sets belong to both the sets in the identity. To include that an element is in a set a 1 is used; to indicate that an element is not in a set, a 0 is used.

Example.

We will use membership table to show that

$$X \cup (Y \cap Z) = (X \cup Y) \cap (X \cup Z).$$

The membership table for these combinations of sets is shown in the following table. This table has eight rows. Since the columns for $X \cup (Y \cap Z)$ and $(X \cup Y) \cap (X \cup Z)$ are the same, the identity is valid.



Membership Tables

	X	Y	Z	$Y \cap Z$	$X \cup (Y \cap Z)$	$X \cup Y$	$X \cup Z$	$(X \cup Y) \cap (X \cup Z)$
1	1	1	1	1	1	1	1	1
2	1	1	0	0	1	1	1	1
3	1	0	1	0	1	1	1	1
4	1	0	0	0	1	1	1	1
5	0	1	1	1	1	1	1	1
6	0	1	0	0	0	1	0	0
7	0	0	1	0	0	0	1	0
8	0	0	0	0	0	0	0	0

Additional set identities can be established using those that we have already proved.

Example.

Let A, B and C be sets. To show that

$$\left(A \cup (B \cap C)\right)^c = (C^c \cup B^c) \cap A^c.$$

We have

$$\begin{aligned}\big(A \cup (B \cap C)\big)^c &= A^c \cap (B \cap C)^c \\ &= A^c \cap (B^c \cup C^c) \\ &= (B^c \cup C^c) \cap A^c \\ &= (C^c \cup B^c) \cap A^c. ■\end{aligned}$$

Since unions and intersections of sets satisfy associative laws, the sets $A \cup B \cup C$ and $A \cap B \cap C$ are well defined when A, B and C are sets. We can also consider unions and intersections of an arbitrary number of sets as follows:

(i) The union of a collection of sets is the set that contain those elements that are members of at least one set in the collection.

We use the notation $A_1 \cup \ldots \cup A_n = \bigcup_{i=1}^{n} A_i$ to denoted the union of the sets $A_1, A_2, \ldots, A_n$

(ii) The intersection of a collection of sets is the set that contains those elements that are members of all the sets in the collection.

We use $A_1 \cap \ldots \cap A_n = \bigcap_{i=1}^{n} A_i$ to denote the intersection of the sets $A_1, A_2, \ldots, A_n$.

●Computer Representation of the sets

There are various ways to represent sets using a computer. We will present a method for storing elements using an arbitrary ordering of the elements of the universal set. This method of representing sets makes computing combinations of sets easy.

Assume that the universal set *U* is finite (and of reasonable size so that the number of elements of *U* is not larger than the memory size of the computer being used). First, specify an arbitrary ordering of the elements of *U*, for instance $a_1, a_2, .., a_n$. Represent a subset *A* of *U* with the bit string of length *n*, where the ith bit in this string is 1 if a_i belongs to *A* and is 0 if a_i does not belong to *A*.

Example.

Let $U = \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10\}$, and the ordering of elements of *U* has the elements in increasing order, i.e, $a_i = i$.

What bit strings represent the subset of all odd integers in *U*, the subset of all even integers in *U*, and the subset of integers not exceeding 5 in *U*?

To do this. The bit string that represents the set of odd integers in *U*, namely, $\{1, 3, 5, 7, 9\}$, has one bit in the first, third, fifth, seventh, and ninth positions, and a zero elsewhere. It is 1010101010.

Similarly, we represent the subset of all even integers in *U*, namely, $\{2, 4, 6, 8, 10\}$ by the string $01\ 0101\ 0101$.

(we have split this bit string of length ten into locks of length four for easy reading since long bit strings are difficult to read).

The set of all integers in U that do not exceed 5, namely, $\{1,2,3,4,5\}$, is represented by the string $11\ 1110\ 0000$.■

Using bit strings to represent sets, it is easy to find complements of sets and unions, intersections, and difference of sets.

To find the bit string for the **complement** of a set from the bit string for that set, we simply change each 1 to 0 and each 0 to 1, since $x \in A$ if and only if $x \notin A^c$.

To obtain the set string for the **union** and **intersection** of two sets we perform bitwise Boolean operations on the bit strings representing the two sets.

The bit in the i$^{\text{th}}$ position of the bit string of the union is 1 if either of the bits in the i$^{\text{th}}$ position in the two strings is 1, and is 0 when both bits are 0.

Hence, the bit string for the **union** is the bitwise **OR** of the bit strings for the two sets.

The bit in the i$^{\text{th}}$ position of the bit string of the **intersection** is 1 when the bits in the corresponding position in the two strings are both 1, and is 0 when either of the two bits is 0 (or both are). Hence, the bit string for the intersection is the bitwise **AND** of the bit strings for the two sets.

Example.

We have seen that the bit string for the set $\{1, 3, 5, 7, 9\}$ (with universal set $\{1, 2, 3, 4, 5, 6, 7, 8, 9, 10\}$) is

$$101010\ 1010$$

The bit string for the complement of this set is obtained by replacing 0's with 1's and vice versa. This yields the string

$$01\ 0101\ 0101$$

which corresponds to the set $\{2, 4, 6, 8, 10\}$.■

Example.

The bit strings for the sets $\{1, 2, 3, 4, 5\}$ and $\{1, 3, 5, 7, 9\}$ are $11\ 1110\ \ 0000$ and $10\ 1010\ \ 1010$, respectively.

We use bit strings to find the union and intersection of these sets.

The bit string for the union of these sets is

$1111100000 \vee 1010101010 = 1111101010,$

which corresponds to the set $\{1, 2, 3, 4, 5, 7, 9\}$.

The bit string for the intersection of these sets is

$1111100000 \wedge 1010101010 = 1010100000,$

which corresponds to the set $\{1, 3, 5\}$.■

Exercise Set (1.1)

1. List the members of these sets

- $\{x: x \text{ is a real number such that } x^2 = 1\}$.
- $\{x: x \text{ is a positive integer less than } 12\}$.
- $\{x: x \text{ is the square of an integer and } x < 100\}$.
- $\{x: x \text{ is an integer such that } x^2 = 2\}$.

2- Use set builder notation to give description of each of these sets.

(a) $\{0, 3, 6, 9, 12\}$;

(b) $\{-3, -2, -1, 0, 1, 2, 3\}$.

3- Determine whether each of these statements is true or false.

(a) $0 \in \phi$; (b) $\phi \in \{0\}$;

(c) $\{0\} \subset \phi$; (d) $\phi \subset \{0\}$;

(e) $\{0\} \in \{0\}$; (f) $\{0\} \subset \{0\}$;

(g) $\{\phi\} \subseteq \{\phi\}$; (h) $\phi \in \{\phi\}$;

(i) $\phi \in \{\phi, \{\phi\}\}$; (i) $x \in \{x\}$;

(k) $\{\phi\} \in \{\{\phi\}\}$; (l) $\{x\} \in \{\{x\}\}$;

(m) $\{x\} \subseteq \{x\}$; (n) $\phi \in \{x\}$.

4- What is the Cartesian product $A \times B$, where A is the set of courses offered by the mathematics department at a university and B is the set of mathematics professors at this university?

5- Let A be a set. Show that $\phi \times A = A \times \phi = \phi$.

6- Find the power set of each of these sets.

(a) $\{a\}$; (b) $\{a, b\}$; (c) $\{\phi, \{\phi\}\}$.

7- Let A be the set of students who live within one mile of school and let B be the set of students who walk to classes.

Describe the students in each of the following sets.

(a) $A \cap B$; (b) $A \cup B$; (c) $A - B$; (d) $B - A$.

8- Let $A = \{1, 2, 3, 4, 5\}$ and $B = \{0, 3, 6\}$. Find

(a) $A \cup B$; (b) $A \cap B$; (c) $A - B$; (d) $B - A$.

9- Let A, B and C be sets. Show that

(a) $(A \cup B) \subseteq (A \cup B \cup C)$;

(b) $(A \cap B \cap C) \subseteq (A \cap B)$;

(c) $(A - B) - C \subseteq A - C$;

(d) $(A - C) \cap (C - B) = \phi$;

(e) $(B - A) \cup (C - A) = (B \cup C) - A$.

10- What can you say about the sets A and B if we know that

(a) $A \cup B = A$;

(b) $A \cap B = A$;

(c) $A - B = A$;

(d) $A \cap B = B \cap A$;

(e) $A - B = B - A$.

11- Suppose that the universal set is $U = \{1,2,3,4,5,6,7,8,9,10\}$. Express each of these sets with bit strings.

(a) $\{3, 4, 5\}$; (b) $\{1, 3, 6, 10\}$; (c)$\{2, 3, 4, 7, 8, 9\}$.

12- Suppose that the universal set is $U = \{1,2,3,4,5,6,7,8,9,10\}$. Find the set specified by each of these bit strings

(a) 11 1100 1111; (b) 01 0111 1000; (c)10 0000 0001.

13- What subsets of a finite universal set do these bit strings represent?

(a) The string with all zeros;

(b) The string with all ones.

14- Let A and B be sets. Show that

(a) $(A \cap B) \subseteq A$;

(b) $A \subseteq (A \cup B)$;

(c) $A - B \subseteq A$;

(d) $A \cap (B - A) = \phi$;

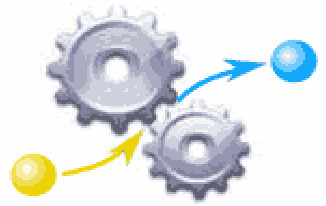
(e) $A \cup (B - A) = A \cup B$.

15- Show that if A, B and C are sets then

$$|A \cup B \cup C| = |A| + |B| + |C| - |A \cap B| - |A \cap C| - |B \cap C| + |A \cap B \cap C|$$

1.2 Functions

What is a Function?



A function relates an input to an output. It is like a machine that has an input and an output. The output is related somehow to the input.

"$f(x) \;= \ldots$" is the classic way of writing a function.

And there are other ways, as you will see!

I will show you many ways to think about functions, but there will always be three main parts:

●The input ● The relationship ● The output

Example.

"Multiply by 2" is a very simple function. Here are the three parts:

Input	***Relationship***	**Output**
0	× 2	0
1	× 2	2
7	× 2	14
10	× 2	20
...	...	...

● Some Examples of Functions

● $\mathbf{x^2}$ (squaring) is a function

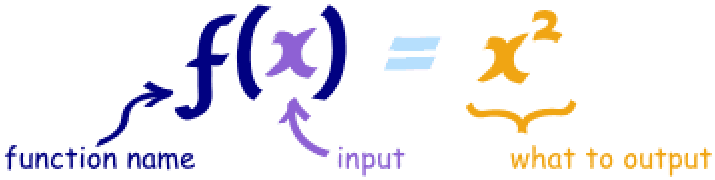
● $\mathbf{x^3+1}$ is also a function

● Sine, Cosine and Tangent are functions used in trigonometry

● Names

First, it is useful to give a function a **name**.

The most common name is "***f***", but you can have other names like "***g***" ....



You would say *"f of x equals x squared"*

The function $f(x) = x^2$ shows you that function "f" takes "x" and squares it.

The "*x*" is Just a Place-Holder!

Don't get too concerned about "*x*", it is just there to show you where the input goes and what happens to it. It could be anything!

So this function: $f(x) = 1 - x + x^2$

Would be the same function if we wrote:

● $f(q) = 1 - q + q^2, h(A) = 1 - A + A^2, w(\theta) = 1 - \theta + \theta^2.$

It is just there so you know where to put the values:

$$f(2) = 1 - 2 + 2^2 = 3$$

Sometimes There is No Function Name

Sometimes a function has no name, and you might just see something like: $y = x^2$. But there is still:

◦ an input (*x*) ◦ a relationship (squaring) ◦ and an output (*y*)

What Types of Things Do Functions Process?

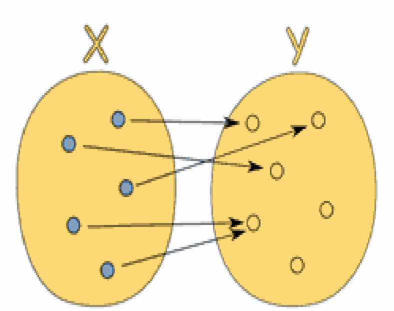
A function takes **elements of a set**, and gives back **elements of a set**.

This can be said in one definition:

Definition.

Let A and B be nonempty sets. A ***function*** f from A to B is an assignment of exactly one element of B to each element of A. We write $f(a) = b$ if b is the unique element of B assigned by the function f to the element a of A. If f is a function from A to B, we write $f : A \to B$.

Remark*:* Functions are sometimes also called **mappings** or **transformations**.



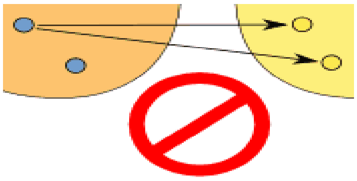
Formal Definition of a Function

A function relates **each element** of a set with **exactly one** element of another set (possibly the same set).

The Two Important Things!

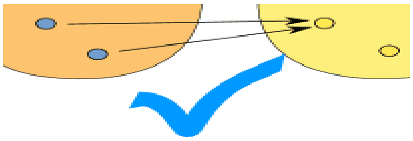
1."...each element..." means that every element in ***X*** is related to some element in ***Y***. We say that the function ***covers X*** (relates every element of it). (But some elements of ***Y*** might not be related to at all, which is fine.)

2. "...exactly one..." means that a function is ***single valued***. It will not give back 2 or more results for the same input.



(one-to-many)

This is **NOT** OK in a function



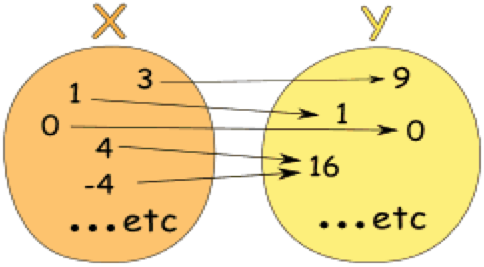
(many-to-one)

But this **is** OK in a function

If a relationship does not follow those two rules then it is **not a function** ... it would still be a relationship, just not a function.

Example

The relationship x → x^2



It is a function, because:

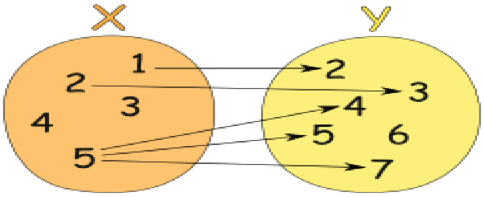
- Every element in X is related to Y.
- No element in X has two or more relationships.

So it follows the rules.

(Notice how both **4** and **-4** relate to **16**, which is allowed.)

Example.

This relationship is not a function:



It is a **relationship**, but it is **not a function**, for these reasons:

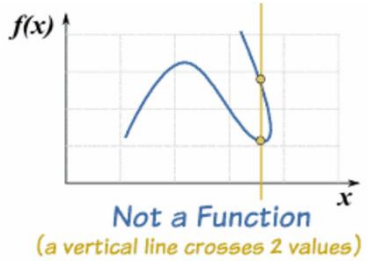
- Value "3" in *X* has no relation in *Y*.
- Value "4" in *X* has no relation in *Y*.
- Value "5" is related to more than one value in *Y*.

(But the fact that "6" in *Y* is not related to does not matter)

●Vertical Line Test

On a graph, the idea of **single valued** means that no vertical line would ever cross more than one value.

If it **crosses more than once** it is still a valid curve, but it would **not be a function**.



□ Set of Ordered Pairs

Here is another way to think about functions:

You can write the input and output of a function as an "ordered pair". They are called **ordered** pairs because the input always comes first, and the output second: (input, output). So it looks like this: $(\,\boldsymbol{x}, \boldsymbol{f}(\boldsymbol{x})\,)$.

Example.

$\{(2,4),(3,5),(7,3)\}$ is a function says "2 is related to 4", "3 is related to 5" and "7 is related 3".

Also, notice that:

But the function has to be **single valued**, so we also say "if it contains (a,b) and (a,c), then b must equal c". Which is just a way of saying that an input of "a" cannot produce two different results.

Example.

$\{(2,4),(2,5),(7,3)\}$ is **not** a function because $(2,4)$ and $(2,5)$ means that 2 could be related to 4 **or** 5. In other words it is not a function because it is **not single valued**

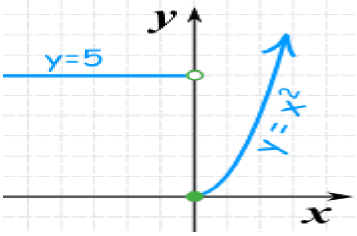
● Piecewise Functions: A Function Can be in Pieces

You can create functions that depending on the input value.

Example

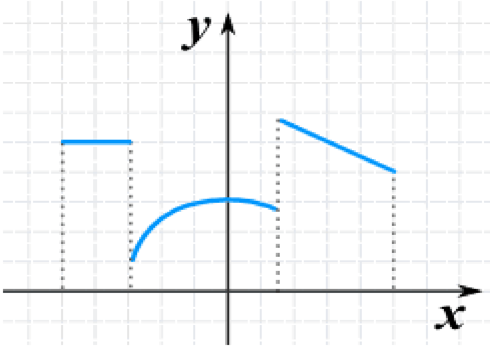
A function with two pieces:

- when *x* is less than 0, it gives 5,
- when *x* is 0 or more it gives x^2.



Example

A function with three pieces:

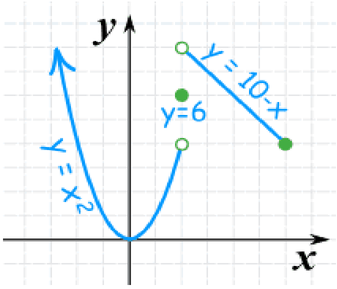


A function made up of 3 pieces

Example

A function with three pieces:

It looks like this:



$$f(x) = \begin{cases} x^2 & x < 2 \\ 6 & x = 2 \\ 10 - x & x > 2 \text{ and } x \leq 6 \end{cases}$$

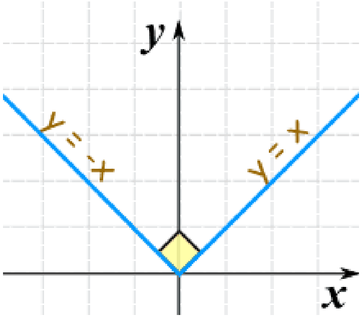
(a solid dot means "including", an open dot means "not including")

● The Absolute Value Function

The Absolute Value Function is a famous Piecewise Function.

It has two pieces:

- below zero: $-\boldsymbol{x}$
- from 0 onwards: $\boldsymbol{x}$
- This is its graph:



$$f(x) = |x| = \begin{cases} x & x \geq 0 \\ -x & x < 0 \end{cases}$$

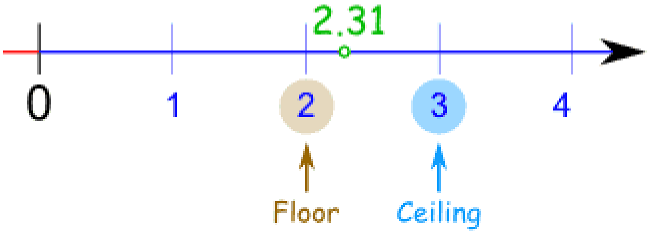
● Floor and Ceiling Functions

The Floor and Ceiling Functions are a Piecewise Functions.

They give you the **nearest integer** up or down.

Example

What is the floor and ceiling of 2.31?



The Floor of 2.31 is **2**

The Ceiling of 2.31 is **3**

What if you want the floor or ceiling of a number that is already an integer? That's easy: no change!

Example.

What is the floor and ceiling of 5?

The Floor of 5 is **5**. The Ceiling of 5 is **5.**

Here are some example values for you:

x	Floor	Ceiling
-1.1	-2	-1
0	0	0
1.01	1	2
2.9	2	3
3	3	3

●Symbols

The symbols for floor and ceiling are like the square brackets [ ] with the top or bottom part missing:

$\lfloor x \rfloor$ floor(x) $\lceil x \rceil$ ceil(x)

But I prefer to use the word form: **floor**(*x*) and **ceil**(*x*)

Definition

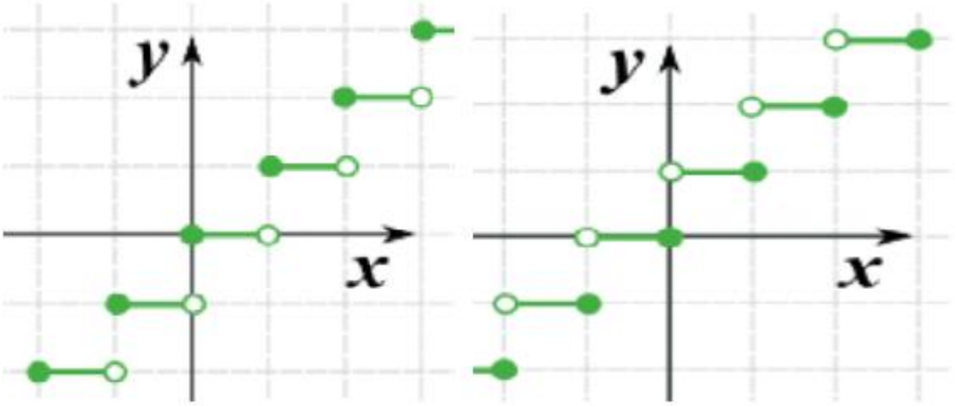
Floor Function:

the greatest integer that is less than or equal to **x.**

Ceiling Function:

the least integer that is greater than or equal to **x.**

● As A Graph



The Floor Function The Ceiling Function

The following table, with x denoting a real number and n is integer, displays some simple but important properties of the floor and ceiling functions.

(1a)	$\lfloor x \rfloor = n$ if and only if $n \leq x < n + 1$
(1b)	$\lceil x \rceil = n$ if and only if $n - 1 < x \leq n$
(1c)	$\lfloor x \rfloor = n$ if and only if $x - 1 < n \leq x$
(1 d)	$\lceil x \rceil = n$ if and only if $x \leq n < x + 1$
(2)	$x - 1 < \lfloor x \rfloor \leq x \leq \lceil x \rceil < x + 1$
(3 a)	$\lfloor -x \rfloor = -\lceil x \rceil$
(3 b)	$\lceil -x \rceil = -\lfloor x \rfloor$
(4 a)	$\lfloor \boldsymbol{x} + \boldsymbol{n} \rfloor = \lfloor \boldsymbol{x} \rfloor + \boldsymbol{n}$
(4 b)	$\lceil \boldsymbol{x} + \boldsymbol{n} \rceil = \lceil \boldsymbol{x} \rceil + \boldsymbol{n}$

Each property in this table can be established using the definitions of the floor and ceiling functions properties (1a), (1b), (1c) and (1d) follow directly from these definitions.

For example (1a) states that $\lfloor x \rfloor = n$ if and only if the integer n is less than or equal to x and $n + 1$ is larger than x. This is precisely what it means for n to be the greatest integer not exceeding x, which is the definition of $\lfloor x \rfloor = n$.

Properties (1b), (1c) and (1d) can be established similarly.

We will prove (4a) as follows:

Suppose that $\lfloor x \rfloor = m$, where m is a positive integer. By (1a) it follows that $m \leq x < m + 1$. Adding n to both sides of this inequality shows that

$$m + n \leq x + n \leq m + n + 1.$$

Using property (1a) again, we see that

$$\lfloor x + n \rfloor = m + n = \lfloor x \rfloor + n.$$

This completes the proof.

Example.

Prove or disprove that $\lceil x + y \rceil = \lceil x \rceil + \lceil y \rceil$ for all real numbers x and y.

Solution.

Although this statement may appear reasonable, it is false.

A counter example is supplied by $x = \frac{1}{2}$ and $y = \frac{1}{2}$. With these values we find that $\lceil x + y \rceil = \left\lceil \frac{1}{2} + \frac{1}{2} \right\rceil = \lceil 1 \rceil = 1$.

But $\lceil x \rceil + \lceil y \rceil = \left\lceil \frac{1}{2} \right\rceil + \left\lceil \frac{1}{2} \right\rceil = 1 + 1 = 2$.

● Domain, Codomain and Range

How to Specify Domains and Ranges

In our examples above

- the set "*X*" is called the **Domain**,
- the set "*Y*" is called the **Codomain**, and
- the set of elements that get pointed to in *Y* (the actual values produced by the function) is called the **Range**.

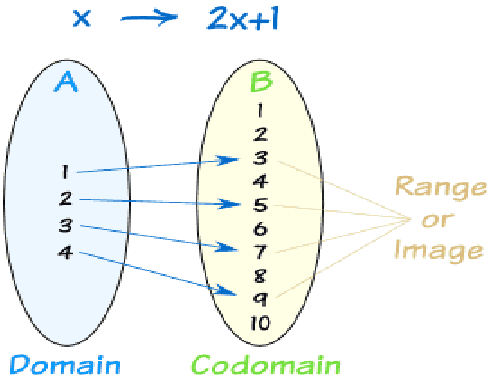
Let us look at a simple example:

In this illustration:

◦the set "*A*" is the Domain,

◦the set "*B*" is the Codomain,

◦and the set of elements that get pointed to in B (the actual values produced by the function) are the Range, also called the Image.



In that example:

- Domain: $\{1, 2, 3, 4\}$
- Codomain: $\{1, 2, 3, 4, 5, 6, 7, 8, 9, 10\}$
- Range: $\{3, 5, 7, 9\}$

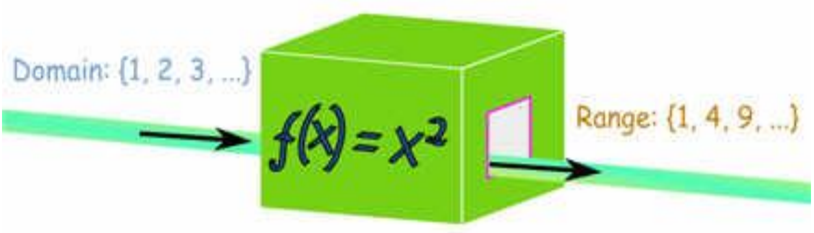
⍰ Part of the Function

Now, what comes **out** *(the Range)* depends on what you put **in** *(the Domain)* ... but **YOU** can define the Domain!

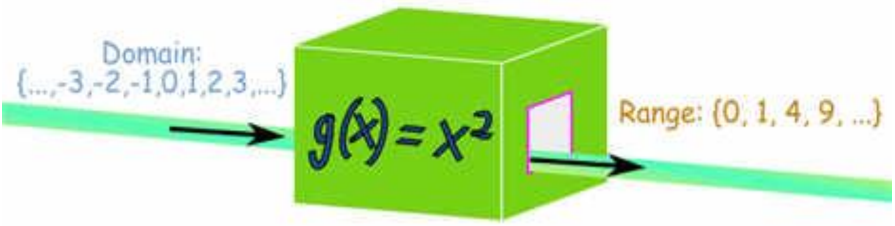
In fact the Domain is an essential part of the function. Change the Domain and you have a different function.

Example.

A simple function like $f(x) = x^2$ can have the **domain** (what goes in) of just the counting numbers $\{1,2,3,...\}$, and the **range** will therefore be the set $\{1,4,9,...\}$



And another function $g(x) = x^2$ can have the domain of integers $\{...,-3,-2,-1,0,1,2,3,...\}$, in which case the range will be the set $\{0,1,4,9,...\}$



Even though both functions take the input and square it, they operate on a **different set of inputs**, and so give a different set of outputs. In this case the range of $g(x)$ also includes 0. Also they will have different properties. For example $f(x)$ always gives a unique answer, but $g(x)$ can give the same answer with two different inputs (such as $g(-2) = 4$, and also $g(2) = 4$). So, the domain is an essential part of the function.

Does Every Function Have a Domain?

Yes, but in simpler mathematics you never notice this, because the domain is ***assumed***:

- Usually it is assumed to be something like "all numbers that would work".
- Or if you are studying whole numbers, the domain is assumed to be whole numbers.

But in more advanced work you need to be more careful!

⍰ Codomain vs Range

The Codomain and Range are both on the output side but are subtly different.

The Codomain is the set of values that could **possibly** come out. The Codomain is actually **part of the definition** of the function. And the Range is the set of values that **actually do** come out.

Example.

you can define a function $\boldsymbol{f(x) = 2x}$ with a domain and codomain of integers (because you say so). But by thinking about it you can see that the range (actual output values) would be just the **even** integers. So the codomain is integers (you defined it that way), but the range is even integers.

The Range is a subset of the Codomain.

Why both? Well, sometimes you don't know the ***exact*** range (because the function may be complicated or not fully known), but you know the set it ***lies in*** (such as integers or reals). So, you define the codomain and continue on.

The Importance of Codomain

Let me ask you a question: Is ***square root*** a function?

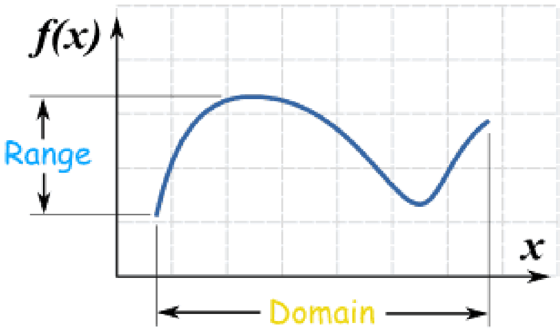
If you say the codomain (the possible outputs) is **the set of real numbers**, then square root is **not a function**! ... is that a surprise? The reason is that there could be two answers for one input, for example $f(9) = 3$ or -3**.**

A function must be ***single valued***. It can not give back 2 or more results for the same input. So "*f*(9) = 3 **or** -3" is not right!

But it can be fixed by simply **limiting the codomain** to non-negative real numbers. In fact, the radical symbol (like $\sqrt{x}$) always means the principal (positive) square root, so $\sqrt{x}$ is a function because its codomain is correct. So, t**hat you choose for the codomain** can actually affect whether something is a **function or not**.

Domains

Now you must consider the **Domains** of the functions.

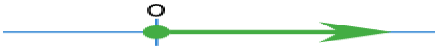


The domain is **the set of all the values** that go into a function.

The function must work for all values you give it, so it is **up to you** to make sure you get the domain correct!

Example: the domain for $\sqrt{x}$ (the square root of x)

You cannot have the square root of a negative number (unless you use imaginary numbers, but we aren't), so we must **exclude** negative numbers: The Domain of $\sqrt{x}$ is all non-negative Real Numbers. On the Number Line it looks like:



Using set-builder notation it is written: $\{x|x \in \mathbb{R}, x \geq 0\}$. Or using interval notation it is: $[0, +\infty)$. It is important to get the Domain right, or you will get bad results! There is also:

Dom(*f*) or Dom *f* meaning "the domain of the function *f*".

Ran(*f*) or Ran *f* meaning "the range of the function *f*".

Definition

If *f* is a function from *A* to *B*, we say that *A* is the *domain* of *f* and *B* is the *codomain* of *f*. If $f(a) = b$, we say that *b* is the *image* of *a* and *a* is a *preimage* of *b*. The *range*, or *image*, of *f* is the set of all images of elements of *A*. Also, if *f* is a function from *A* to *B*, we say that *f maps A* to *B*.

Example

Let *f* be the function that assigns the last two bits of a bit string of length 2 or greater to that string. For example, $f(11010) = 10$. Then, the domain of *f* is the set of all bit strings of length 2 or greater, and both the codomain and range are the set $\{00, 01, 10, 11\}$.■

Example.

Let $f: \mathbb{Z} \to \mathbb{Z}$ assign the square of an integer to this integer. Then, $f(x) = x^2$, where the domain of f is the set of all integers, the codomain of f is the set of all integers, and the range of f is the set of all integers that are perfect squares, namely, $\{0, 1, 4, 9, \ldots\}$.■

Example.

The domain and codomain of functions are often specified in programming languages. For instance, the Java statement

int **floor**(float real){. . .}

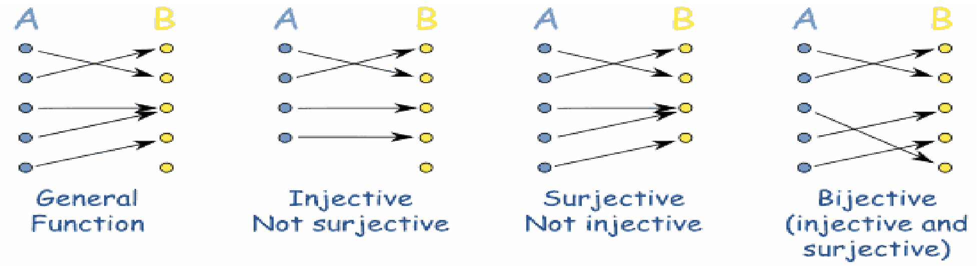
and the C++ function statement

int **function** (float *x*){. . .}

both tell us that the domain of the floor function is the set of real numbers (represented by floating point numbers) and its codomain is the set of integers. ■

A function is called **real-valued** if its codomain is the set of real numbers, and it is called **integer-valued** if its codomain is the set of integers.

● Injective, Surjective and Bijective



A **General Function** points from each member of "*A*" to a member of "*B*".

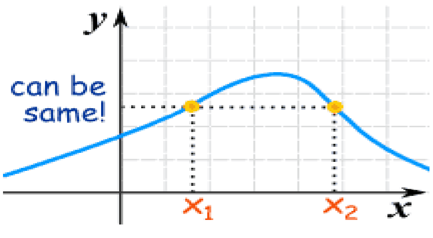
Injective means that every member of "*A*" has **its own unique** matching member in "*B*". As it is also a function **one-to-many is not OK.** And you won't get two "*A*"s pointing to the same "*B*", so **many-to-one is NOT OK**. But you can have a "*B*" without a matching "*A*". Injective functions can be **reversed**!

If "*A*" goes to a unique "*B*" then given that "*B*" value you can go back again to "*A*" (this would not work if two or more "*A*"s pointed to one "*B*" like in the "General Function"). Injective is also called "**one-to-one**".

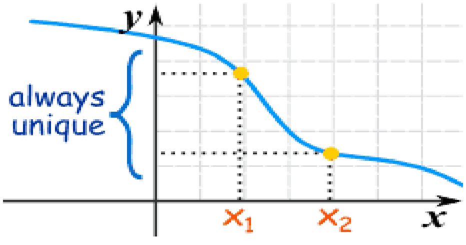
Surjective means that every "*B*" has **at least one** matching "*A*" (maybe more than one). There won't be a "*B*" left out. **Bijective** means both Injective and Surjective together. So there is a perfect "**one-to-one correspondence**" between the members of the sets. (But don't get that confused with the term "one-to-one" used to mean injective).

□ On The Graph

Let me show you on a graph what a "**General Function**" and a "**Injective Function**" looks like:



General Function



"Injective" (one-to-one)

In fact you can do a "Horizontal Line Test":

To be **Injective**, a Horizontal Line should never intersect the curve at 2 or more points. *Note that: Strictly Increasing (and Strictly Decreasing) functions are Injective.*

● Formal Definitions

A function *f* is said to be *one-to-one* if and only if $f(a) = f(b)$ implies that $a = b$ for all *a* and *b* in the domain of *f*. A function is said to be *injective* if it is one-to-one.

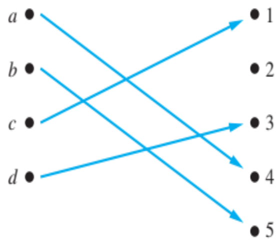
We illustrate this concept by giving examples of functions that are one-to-one and other functions that are not one-to-one.

Example.

Determine whether the function *f* from $\{a, b, c, d\}$ to $\{1, 2, 3, 4, 5\}$ with $f(a) = 4$, $f(b) = 5$, $f(c) = 1$, and $f(d) = 3$ is one-to-one.

Solution.

The function *f* is one-to-one because *f* takes on different values at the four elements of its domain. This is illustrated in the figure. ■



Example.

$f(x) = x + 5$ from the set of real numbers ℝ to ℝ is an injective function. This function can be easily reversed. for example: $f(3) = 8$. Given 8 we can go back to 3. ■

Example.

$f(x) = x^2$ from the set of real numbers $\mathbb{R}$ to $\mathbb{R}$ is not an injective function because: $f(2) = 4$ and $f(-2) = 4$. This is against the definition $f(x) = f(y), x = y$, because $f(2) = f(-2)$ but $2 \neq -2$. In other words, there are **two** values of "A" that point to one "B", and this function could not be reversed (given the value "4" ... what produced it?). BUT if we made it from the set of natural numbers $\mathbb{N}$ to $\mathbb{N}$ then it is injective, because: $f(2) = 4$. There is no $f(-2)$, because -2 is not a natural number. ■

Example.

Study the injection of the function $f(x) = ax + b$.

Solution.

Let $f(x_1) = f(x_2)$ then $ax_1 + b = ax_2 + b$ Thus, $ax_1 = ax_2$. If $a \neq 0$,then $x_1 = x_2$ and f is one-to-one. If $a = 0$, then $f(x) = b$ for every $x \in \mathbb{R}$ and $f(1) = f(2)$, for example. Hence f is not one-to-one. ■

Example.

The function $f : \mathbb{R} \to \mathbb{R}$, defined by $f(x) = x^2$ is not one-to-one. For example, $f(-3) = f(3) = 9$. But, if we restrict the function on the interval $[0, \infty)$, the function will be one-to-one as $f(x_1) = f(x_2) \Rightarrow x_1^2 = x_2^2 \Rightarrow x_1 = \pm x_2$. Since x_1, x_2 are positive then $x_1 = x_2$. ■

Example.

Prove that $f(x) = \frac{3x+1}{5-2x}$ is injective.

Solution.

Let $x_1, x_2 \in D(f) = \mathbb{R} - \{5/2\}$, then

$$f(x_1) = f(x_2) \Rightarrow \frac{3x_1+1}{5-2x_1} = \frac{3x_2+1}{5-2x_2}$$

$$\Rightarrow 15x_1 - 6x_1 x_2 + 5 - 2x_2 = 15x_2 - 6x_1x_2 + 5 - 2x_1 \Rightarrow$$

$$x_1 = x_2.$$

Thus, f is injective.■

● Surjective (Also Called "Onto")

A function *f* (from set *A* to *B*) is **surjective** if and only for every *y* in *B*, there is at least one *x* in *A* such that f(x) = y. In other words *f* is surjective if and only if $f(A) = B$. So, every element of the range corresponds to at least one member of the domain.

Remark

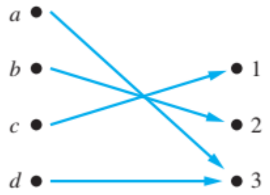
A function *f* is onto if $\forall y \exists x (f(x) = y)$, where the domain for *x* is the domain of the function and the domain for *y* is the codomain of the function.

Example

Let *f* be the function from $\{a, b, c, d\}$ to $\{1, 2, 3\}$ defined by $f(a) = 3, f(b) = 2, f(c) = 1$, and $f(d) = 3$.

Is *f* an onto function?

Solution



Because all three elements of the codomain are images of elements in the domain, we see that *f* is onto. This is illustrated in the figure. Note that if the codomain were $\{1, 2, 3, 4\}$, then *f* would not be onto. ■

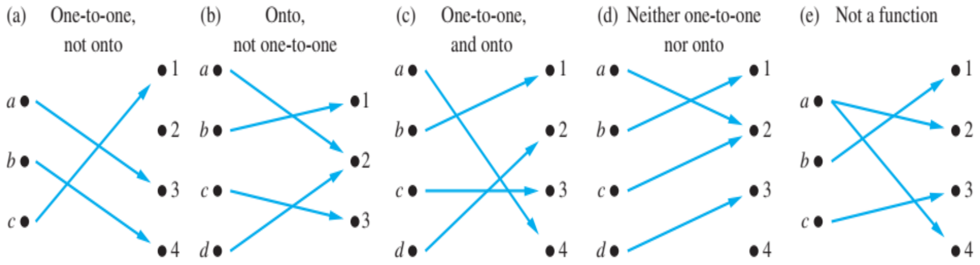
Example

The function $f(x) = 2x$ from the set of natural numbers $\mathbb{N}$ to the set of non-negative even numbers is a surjective function. However, $f(x) = 2x$ from the set of natural numbers $\mathbb{N}$ to $\mathbb{N}$ is not surjective, because, for example, nothing in $\mathbb{N}$ can be mapped to 3 by this function. ■

Example

The function $f: \mathbb{R} \to \mathbb{R}$ defined by $f(x) = x^2$ is not surjective because $\mathrm{Im}(f) = [0\,, \infty) \neq \mathbb{R}$. But, the function $g: \mathbb{R} \to \mathbb{R}$ defined by $f(x) = x^3$ is surjective because $\mathrm{Im}(f) = \mathbb{R}$. ■

Examples of Different Types of Correspondences.



(a) One-to-one, not onto

(b) Onto, not one-to-one

(c) One-to-one, and onto

(d) Neither one-to-one nor onto

(e) Not a function

● Bijective

A function f (from set A to B) is bijective if, for every y in B, there is exactly one x in A such that f(x) = y. Alternatively, f is bijective if it is a one-to-one correspondence between those sets, in other words both injective and surjective.

Example

The function $f(x) = x^2$ from the set of positive real numbers to positive real numbers is injective and surjective. Thus, it is also bijective. But not from the set of real numbers ℝ because you could have, for example, both $f(2) = 4$ and $f(-2) = 4$. ■

Exercise. Which of the following functions is NOT injective?

A) $f(x) = x^3 + 4$ from ℝ to ℝ; B) $f(x) = x^3 + 4$ from ℕ to ℕ

C) $f(x) = x^2 + 4$ from ℝ to ℝ; D) $f(x) = x^2 + 4$ from ℕ to ℕ

● Inverse function

Let f be a one-to-one correspondence from the set A to the set B. The inverse function of f is the function that assigns to an element b belonging to B the unique element a in A such that $f(a) = b$. The inverse function of f is denoted by f^{-1}. Hence $f^{-1}(b) = a$ when $f(a) = b$. If a function f is not a one-to-one correspondence, we cannot define an inverse function of f. When f is not one-to-one correspondence, either it is not one-to-one, or it is not onto. If f is not one-to-one, some element b in the codomain is the image of more than one element in the domain. If *f* is not onto, for some element b in the codomain, no element a in the domain exists for which $f(a) = b$. Consequently, if *f* is not a 1-1 correspondence, we cannot assign to each element b in the codomain a unique element a in the domain such that $f(a) = b$.

Example

Let f be the function from $\{a, b, c\}$ to $\{1,2,3\}$ such that $f(a) = 2, f(b) = 3$, and $f(c) = 1$. The function *f* is invertible since it is one-to-one correspondence. The inverse function f^{-1} reverses the correspondence given by *f*, so that $f^{-1}(1) = c, f^{-1}(2) = a$ and $f^{-1}(3) = b$. ■

Example

The function $f\colon \mathbb{Z} \to \mathbb{Z}$such that $f(x) = x + 1$ has an inverse since it is a one-to-one correspondence. To reverse the correspondence, suppose *y* is the image of *x*, so that $y = x + 1$. Then $x = y - 1$. This means that $y - 1$ is the unique element of $\mathbb{Z}$ that is sent to *y* by *f*. Consequently, $f^{-1}(y) = y - 1$. ■

Example

The function $f\colon \mathbb{Z} \to \mathbb{Z}$ with $f(x) = x^2$ is not invertible since *f* is not one-to-one, since, for instance, $f(1) = f(-1) = 1$. ■

● Operations with Functions

You can add, subtract, multiply and divide functions!

The result will be a new function

Let us try doing those operations on $f(x)$ and $g(x)$.

Addition

You can add two functions: $(f+g)(x) = f(x) + g(x)$

Note:*put the $f+g$ inside ( ) so you know they both work on x.*

Example

Let $\boldsymbol{f(x) = 2x+3}$ and $\boldsymbol{g(x) = x^2}$. Then, we have

$\boldsymbol{(f+g)(x) = (2x+3) + (x^2) = x^2+2x+3}$. ■

Example

Let $\boldsymbol{v(x) = 5x+1}$ and $\boldsymbol{w(x) = 3x-2}$. Then, we have

$\boldsymbol{(v+w)(x) = (5x+1) + (3x-2) = 8x-1}$. ■

The only other thing to worry about is the Domain (the set of numbers that go into the function), but I will talk about that later!

Subtraction

You can subtract two functions:

$(f-g)(x) = f(x) - g(x)$.

Example

$f(x) = 2x+3$ and $g(x) = x^2$. Then, we have

$(f-g)(x) = (2x+3) - (x^2)$. ■

Multiplication

You can multiply two functions:

$$(f \cdot g)(x) = f(x) \cdot g(x)$$

Example

$f(x) = 2x + 3$ and $g(x) = x^2$,

$(f \cdot g)(x) = (2x + 3)\,(x^2) = 2x^3 + 3x^2$. ■

Division

And you can divide two functions:

$$(f/g)(x) = f(x)\,/\,g(x).$$

Example

$f(x) = 2x + 3$ and $g(x) = x^2$,

$(f/g)(x) = (2x + 3)\,/(x^2)$. ■

How to Work Out the New Domain

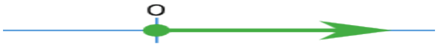
When you do operations on functions, you end up with the **restrictions of both**.

It is like cooking for friends: one can't eat peanuts, the other can't eat dairy food. So what you cook can't have peanuts **and also** can't have dairy products.

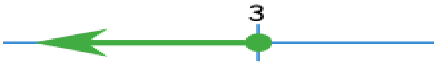
Example

$f(x) = \sqrt{x}$ and $g(x) = \sqrt{(3 - x)}$

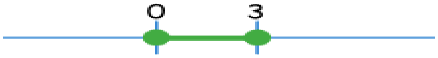
The domain for $f(x) = \sqrt{x}$ is from 0 onwards:



The domain for $g(x) = \sqrt{(3 - x)}$ is up to and including 3:



The new domain (after adding or whatever) is therefore from 0 to 3:



If you choose any other value, then one or the other part of the new function won't work. In other words, you want to find where the two domains **intersect**.■

Note

We can put this whole idea into one line using Set Builder Notation:

$\mathrm{Dom}(f+g)=\{x\in\mathbb{R}: x\in\mathrm{Dom}(f) \text{ and } x\in\mathrm{Dom}(g)\}.$

Which says "the domain of f plus g is the set of all Real Numbers that are in the domain of f AND in the domain of g"

The same rule applies when you add, subtract, multiply or divide, except divide has one extra rule.

□ An Extra Rule for Division

There is an **extra rule** for division:

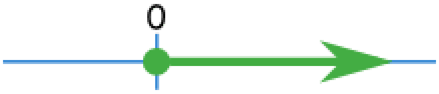
As well as restricting the domain as above, when we **divide**:

$$(f/g)(x) = f(x)\,/\,g(x)$$

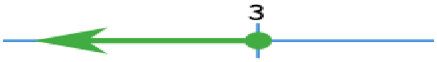
we must also make sure that $g(x)$ is **not equal to zero** (so we don't divide by zero).

Example: $\boldsymbol{f(x)=\sqrt{x}}$ and $\boldsymbol{g(x)=\sqrt{3-x}}$. $\boldsymbol{\left({}^{f}/_{g}\right)(x)=\frac{\sqrt{x}}{\sqrt{3-x}}}$.

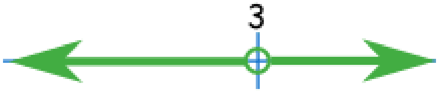
The domain for **:** $\boldsymbol{f(x)=\sqrt{x}}$ is from 0 onwards:



The domain for $\boldsymbol{g}(\boldsymbol{x}) = \sqrt{\boldsymbol{3}-\boldsymbol{x}}$ is up to and including 3:

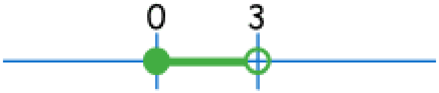


But we also have the restriction that $\sqrt{\boldsymbol{3}-\boldsymbol{x}}$ **cannot be zero**, so x cannot be 3:



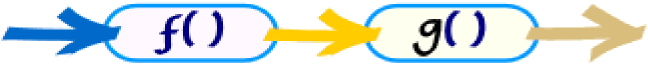
(Notice the **open circle** at 3, which means **not including** 3)

So all together we end up with:



● Composition of Functions

"Function Composition" is applying one function to the results of another:



The result of $f()$ is sent through $g()$

It is written: $(g \circ f)(x)$. Which means: $g(f(x))$.

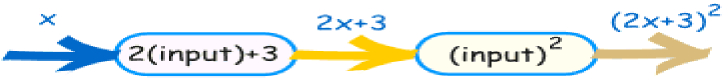
Example

$f(x) = 2x + 3$ and $g(x) = x^2$. "x" is just a placeholder, and to avoid confusion let's just call it "input":

$f(\text{input}) = 2(\text{input}) + 3,\ g(\text{input}) = (\text{input})^2.$

So, let's start: $(g \circ f)(x) = g(f(x))$.

First we apply *f*, then apply *g* to that result:

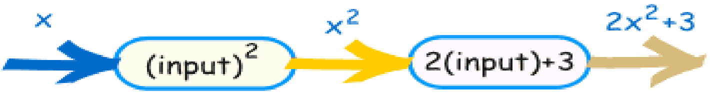


$$(g \circ f)(x) = (2x+3)^2$$

What if we reverse the order of *f* and *g*?

$$(f \circ g)(x) = f\big(g(x)\big).$$

First we apply *g*, then apply *f* to that result:



$$(f \circ g)(x) = 2x^2 + 3$$

We got a different result! So be careful which function comes first.■

● Symbol

The symbol for composition is a small circle: $(g \circ f)(x)$.

It is **not** a filled in dot: $(g.f)(x)$ as that would mean **multiply**.

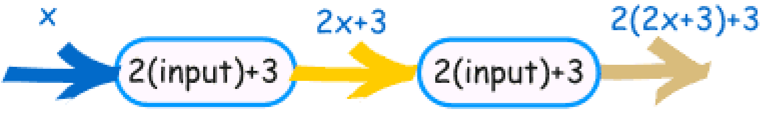
● Composed With Itself

You can even compose a function with itself!

Example

$$f(x) = 2x + 3. \ (f \circ f)(x) = f(f(x)).$$

First we apply *f*, then apply *f* to that result:



$$(f \circ f)(x) = 2(2x+3) + 3 = 4x + 9$$

You should be able to do this without the pretty diagram:

$$\begin{aligned}(f \circ f)(x) &= f\big(f(x)\big)\\ &= f(2x+3)\\ &= 2(2x+3)+3\\ &= 4x + 9.\end{aligned}$$ ■

● Domain of Composite Function

You must get **both Domains** right (the composed function **and** the first function used). When doing, for example, $(g \circ f)(x) = g(f(x))$:

- Make sure you get the Domain for $f(x)$ right,
- Then also make sure that $g(x)$ gets the correct Domain.

Example

$\boldsymbol{f(x) = \sqrt{x}}$ and $\boldsymbol{g(x) = x^2}$. The Domain of $\boldsymbol{f(x) = \sqrt{x}}$ is all non-negative Real Numbers. The Domain of $\boldsymbol{g(x) = x^2}$ is all the Real Numbers. The composed function is:

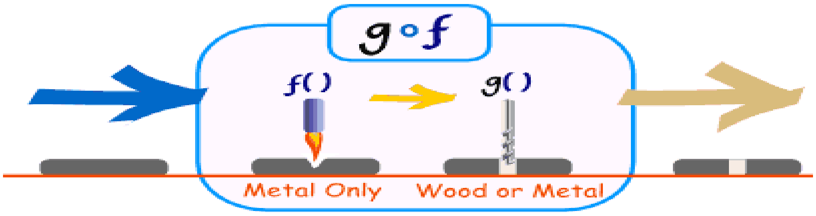
$$(g \circ f)(x) = g(f(x)) = (\sqrt{x})^2 = x$$

Now, "x" would normally have the Domain of all Real Numbers ...... but because it is a **composed function** you must **also consider** $f(x)$.

So the Domain is all non-negative Real Numbers. ■

Why Both Domains?

Well, imagine the functions were machines ... the first one melts a hole with a flame (only for metal), the second one drills the hole a little bigger (works on wood or metal):



> What you see at the end is a drilled hole, and you may think "that should work for wood **or** metal". But if you put wood into $g \circ f$ then the first function f would make a fire and burn everything down!

So what happens "inside the machine" is important.

Example.

Let $f(x) = x^2$ and $g(x) = 3x + 5$.

Find $(f \circ g)(x)$, $(g \circ f)(x)$ Dom $(f \circ g)$ and Dom $(g \circ f)$.

Solution.

$(f \circ g)(x) = f\big(g(x)\big) = f(2x + 3) = (2x + 3)^2$

$(g \circ f)(x) = g\big(f(x)\big) = g(x^2) = 2x^2 + 3$

It is obvious that $\text{Dom}(f) = \mathbb{R}$ and $\text{Dom}(g) = \mathbb{R}$. So,

$\text{Dom}\ (f \circ g) = \{x : x \in \text{Dom}(g), g(x) \in \text{Dom}(f)\}$

$= \{x : x \in \mathbb{R} , 2x + 3 \in \mathbb{R}\} = \mathbb{R}$

$\text{Dom}(g \circ f) = \{x : x \in \text{Dom}(f), f(x) \in \text{Dom}(g)\}$

$= \{x : x \in \mathbb{R} , x^2 \in \mathbb{R}] = \mathbb{R}.$ ■

Example.

Let $f(x) = \frac{x}{x+2}$, $g(x) = \frac{x-1}{x}$.

Find $(f \circ g)(x)$, $(g \circ f)(x)$, $\text{Dom}(f \circ g)$ and $\text{Dom}(g \circ f)$.

Solution.

Note that $\text{Dom}(f) = \mathbb{R} - \{2\}$, and $\text{Dom}(g) = \mathbb{R} - \{0\}$

$$(f \circ g)(x) = f\big(g(x)\big) = f\left(\frac{x-1}{x}\right) = \frac{\frac{x-1}{x}}{\frac{x-1}{x}+2} = \frac{x-1}{3x-1}$$

$$(g \circ f)(x) = g\big(f(x)\big) = g\left(\frac{x}{x+2}\right) = \frac{\frac{x}{x+2}-1}{\frac{x}{x+2}} = \frac{-2}{x}$$

$$\begin{aligned}\text{Dom}(f \circ g)(x) &= \{x : x \in \text{Dom}(g), g(x) \in \text{Dom}(f)\}\\ &= \left\{x: x \neq 0 , \tfrac{x-1}{x} \neq -2\right\}\\ &= \left\{x: x \neq 0 , x \neq \tfrac{1}{3}\right\} = \mathbb{R} - \left\{0 , \tfrac{1}{3}\right\}.\end{aligned}$$

$$\begin{aligned}\text{Dom}(g \circ f)(x) &= \{x : x \in \text{Dom}(f), f(x) \in \text{Dom}(g)\}\\ &= \left\{x : x \neq -2 , \frac{x}{x+2} \neq 0\right\}\\ &= \{x : x \neq -2 , x \neq 0\} = \mathbb{R} - \{0 , -2\}.\end{aligned}$$ ■

Remark

(i) Composition of function is not abelian (commutative), i.e,

$(f \circ g)(x) \neq (g \circ f)(x)$. In general, $\text{Dom}(f \circ g) \neq \text{Dom}(g \circ f)$

(ii) Composition of function is associative:

$$\begin{aligned}\text{Dom}\big(f \circ (g \circ h)\big) &= \text{Dom}\big((f \circ g) \circ h\big)\\ &= \{x : x \in D(h), h(x) \in \text{Dom}(g), g(h(x)) \in \text{Dom}(f)\}.\end{aligned}$$

Also, $\big(f \circ (g \circ h)\big)(x) = (f \circ g) \circ h(x) = f(g(h(x)))$.

(iii) In the above example, we note that:

$\mathrm{Dom}(g \circ f) = \mathbb{R} - \{0\,, -2\}$ and $(g \circ f)(x) = -\frac{2}{x}$ and $\mathrm{Dom}\left(-\frac{2}{x}\right) = \mathbb{R} - \{0\}$. So, we cannot find the domain of the composite function by, addition, subtraction, product, quotient, or composition by final rule.

● De-Composing Function

You can go the other way and break up a function into a composition of other functions. For example: $(x + 1/x)^2$.

That function could have been made from these two functions: $f(x) = x + 1/x,\ g(x) = x^2$. And we would have:

$$(g \circ f)(x) = g(f(x)) = g(x + 1/x) = (x + 1/x)^2$$

This can be useful if the original function is too complicated to work on. Note that the composition $f \circ g$ cannot be defined unless the range of *g* is a subset of the domain of *f*. Also, note that even though $f \circ g$ and $g \circ f$ are defined for the functions *f* and *g*. But $f \circ g$ and $g \circ f$ are not equal. In other words, the commutative law does not hold for the composition of functions.

Remark.

Suppose that $f\colon A \to B$ is a one-to-one correspondence. Then the inverse function $f^{-1}\colon B \to A$ exists and a one-to-one correspondence. f^{-1} reverse the correspondence of *f*, so that $f^{-1}(b) = a$ when *f*(*a*) = *b* and *f*(*a*) = *b* when $f^{-1}(b) = a$. Hence $(f^{-1} \circ f)(a) = f^{-1}(f(a)) = f^{-1}(b) = a$ and $(f \circ f^{-1})(b) = f(f^{-1}(b)) = f(a) = b$. Consequently, $f^{-1} \circ f = I_A$ and $f \circ$

$f^{-1} = I_B$, where I_A and I_B are the identity function on sets A and B, respectively. That is $(f^{-1})^{-1} = f$.

● Cardinality

We defined the cardinality of a finite set as the number of elements in the set. We use the cardinalities of finite sets to tell us when they have the same size, or when one is bigger than the other. In this section we extend this notion to infinite sets. We will be particularly interested in countably infinite sets, which are sets with the same cardinality as the set of positive integers.

The concepts developed in this section have important applications to computer science. A function is called uncomputable if no computer program can be written to find all its values, even with unlimited time and memory.

Definition

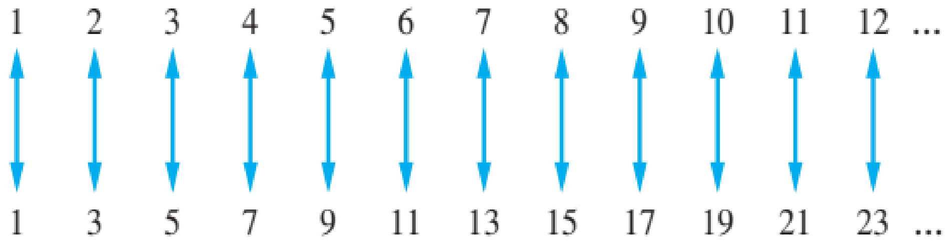
The sets A and B have the same **cardinality** if and only if there is a one-to-one correspondence from A to B. When A and B have the same cardinality, we write $|A| = |B|$.

Definition

If there is a one-to-one function from A to B, the cardinality of A is less than or the same as the cardinality of B and we write $|A| \leq |B|$. Moreover, when $|A| \leq |B|$ and A and B have different cardinality, we say that the cardinality of A is less than the cardinality of B and we write $|A| < |B|$.

● Countable Set

We will now split infinite sets into two groups, those with the same cardinality as the set of natural numbers and those with a different cardinality. The following graph shows a one-to-one correspondence between $\mathbb{Z}^+$ and the set of odd positive integers.



Definition

A set that is either finite or has the same cardinality as the set of positive integers is called **countable**. A set that is not countable is called **uncountable**. When an infinite set S is countable, we denote the cardinality of S by $\aleph_\circ$ (where $\aleph$ is aleph, the first letter of the Hebrew alphabet). We write $|S| = \aleph_\circ$ and say that S has cardinality "aleph null".

Example.

Show that the set of odd positive integers is countable?

Solution

To show that the set of odd positive integers is countable, we will exhibit a one-to-one correspondence between this set and the set of positive integers.

Consider the function

$$f(n) = 2n - 1$$

from $\mathbb{Z}^{+}$ to the set of odd positive integers. We show that f is a one-to-one correspondence by showing that it is both one-to-one and onto. To see that it is one-to-one, suppose that $f(n) = f(m)$. Then $2n - 1 = 2m - 1$, so $n = m$. To see that it is onto, suppose that o is an odd positive integer. Then o is 1 less than an even integer $2k$, where k is a natural number. Hence $o = 2k = f(k)$. We displayed this one-to-one correspondence in the above figure.■

Example

Show that the set of all integers is countable.?

Solution

To show that the set of all integers is countable. we can list all integers in a sequence by starting with 0 and alternating between positive and negative integers; $0, 1, -1, 2, -2, \ldots$. Alternately, we could find a one-to-one correspondence between the set of positive integers and the set of all integers. We leave it to the reader to show that the function $f(n) = n/2$ when n is even and $f(n) = -(n - 1)/2$ when n is odd is such a function. Consequently, the set of all integers is countable.■

It is not surprising that the set of odd integers and the set of all integers are both countable sets. Now we show that the set of rational numbers also is countable.

Example

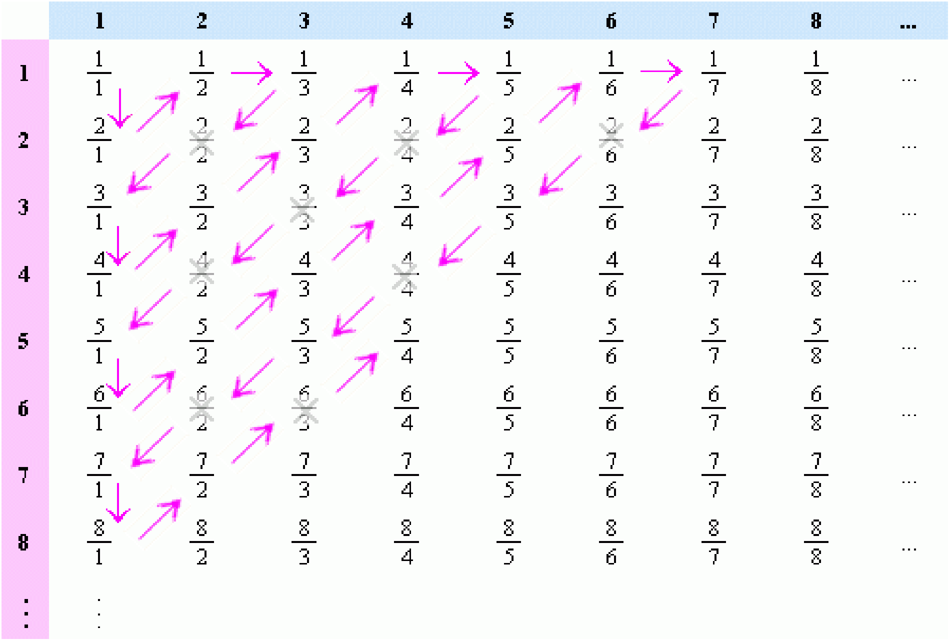
Show that the set of positive rational numbers is countable.?

Solution

We can list the positive rational numbers as a sequence $r_1, r_2, \ldots, r_n, \ldots$.

First note that every positive rational number is the quotient p/q of two positive integers. We can arrange the positive rational numbers by listing those with denominator $q = 1$ in the first row, those with denominator $q = 2$ in the second row, and so on, as displayed in the following figure.

The key to listing the rational numbers in a sequence is to first list the positive rational numbers p/q with $p + q = 2$, followed by those with $p + q = 3$, followed by those with $p + q = 4$, and so on, following the path shown in below figure. Whenever we encounter a number p/q that is already listed, we do not list it again. For example, when we come to $2/2 = 1$ we do not list it because we have already listed $1/1 = 1$. The initial terms in the list of positive rational numbers we have constructed are $1, 1/2$, 2,3, $1/3$, $1/4, 2/3$, $3/2$, 4,5, and so on. These numbers are not deleted; the other numbers in the list are those we leave out because they are already listed. Because all positive rational numbers are listed once, as the reader can verify, we have shown that the set of positive rational numbers is countable.■



	1	2	3	4	5	6	7	8	...
1	$\frac{1}{1}$	$\frac{1}{2}$	$\frac{1}{3}$	$\frac{1}{4}$	$\frac{1}{5}$	$\frac{1}{6}$	$\frac{1}{7}$	$\frac{1}{8}$	...
2	$\frac{2}{1}$	$\cancel{\frac{2}{2}}$	$\frac{2}{3}$	$\cancel{\frac{2}{4}}$	$\frac{2}{5}$	$\cancel{\frac{2}{6}}$	$\frac{2}{7}$	$\frac{2}{8}$	...
3	$\frac{3}{1}$	$\frac{3}{2}$	$\cancel{\frac{3}{3}}$	$\frac{3}{4}$	$\frac{3}{5}$	$\frac{3}{6}$	$\frac{3}{7}$	$\frac{3}{8}$	...
4	$\frac{4}{1}$	$\cancel{\frac{4}{2}}$	$\frac{4}{3}$	$\cancel{\frac{4}{4}}$	$\frac{4}{5}$	$\frac{4}{6}$	$\frac{4}{7}$	$\frac{4}{8}$	...
5	$\frac{5}{1}$	$\frac{5}{2}$	$\frac{5}{3}$	$\frac{5}{4}$	$\frac{5}{5}$	$\frac{5}{6}$	$\frac{5}{7}$	$\frac{5}{8}$	...
6	$\frac{6}{1}$	$\cancel{\frac{6}{2}}$	$\cancel{\frac{6}{3}}$	$\frac{6}{4}$	$\frac{6}{5}$	$\frac{6}{6}$	$\frac{6}{7}$	$\frac{6}{8}$	...
7	$\frac{7}{1}$	$\frac{7}{2}$	$\frac{7}{3}$	$\frac{7}{4}$	$\frac{7}{5}$	$\frac{7}{6}$	$\frac{7}{7}$	$\frac{7}{8}$	...
8	$\frac{8}{1}$	$\frac{8}{2}$	$\frac{8}{3}$	$\frac{8}{4}$	$\frac{8}{5}$	$\frac{8}{6}$	$\frac{8}{7}$	$\frac{8}{8}$	...
⋮	⋮								

Example.

Prove that the set of real numbers is not countable.

Solution

To show that the set of real numbers is uncountable, we suppose that the set of real numbers is countable and arrive at a contradiction. Then, the subset of all real numbers that fall between 0 and 1 would also be countable (any subset of a countable set is also countable). Under this assumption, the real numbers between 0 and 1 can be listed in some order, say, $r_1, r_2, r_3, \ldots$. Let the decimal representation of these real numbers be

$$r_1 = 0.d_{11}d_{12}d_{13}d_{14}\ldots$$

$$r_1 = 0.d_{11}d_{12}d_{13}d_{14}\ldots$$

$$r_3 = 0.d_{31}d_{32}d_{33}d_{34}\ldots$$

$$r_4 = 0.d_{41}d_{42}d_{43}d_{44}\dots$$

$\vdots$

where $d_{ij} \in$ {0,1,2,3,4,5,6,7,8,9}. (For example, if $r_1 =$ $0.23794102\dots$, we have $d_{11} = 2, d_{12} = 3, d_{13} = 7,$ and so on). Then, form a new real number with decimal expansion $r =$ $0.d_1d_2d_3d_4\dots$, where

$$d_i = \begin{cases} 4 & if d_{ii} \neq 4 \\ 5 & if d_{ii} = 4 \end{cases}$$

(As an example, suppose that $r_1 = 0.23794102\dots$, $r_2 =$ $0.44590138\dots$, $r_3 = 0.09118764\dots$, $r_4 = 0.80553900\dots$, and so on. Then we have $r = 0.d_1d_2d_3d_4\dots = 0.4544\dots$, where $d_1 = 4$ because $d_{11} \neq 4$, $d_2 = 5$ because $d_{22} = 4$, $d_3 = 4$ because $d_{33} \neq$ 4, $d_4 = 4$because $d_{44} \neq 4$, and so on).

Every real number has a unique decimal expansion. Then the real number r is not equal to any of $r_1, r_2, r_3, \dots$ because the decimal expansion of r differs from the decimal expansion of r_i in the ith place to the right of the decimal point, for each i.

Because there is a real number r between 0 and 1that is not in the list, the assumption that all the real numbers between 0 and 1 cannot be listed, so the set of real numbers between 0 and 1 is uncountable. Any set with an uncountable subset is uncountable. Hence, the set of real numbers is uncountable.■

Theorem

If A and B are countable sets, then $A \cup B$ is also countable.

Theorem (SCHRÖDER-BERNSTEIN THEOREM)

If A and B are sets with $|A| \leq |B|$ and $|B| \leq |A|$, then $|A| = |B|$. In other words, if there are one-to-one functions f from A to B and g from B to A, then there is a one-to-one correspondence between A and B.

Example.

Show that the $|(0,1)| = |(0,1]|$.

Solution

It is not at all obvious how to find a one-to-one correspondence between $(0,1)$ and $(0,1]$ to show that $|(0,1)| = |(0,1]|$. Fortunately, we can use the Schröder-Bernstein theorem instead. Finding a one-to-one function from $(0,1)$ to $(0,1]$ is simple. Because $(0,1) \subset (0,1]$, $f(x) = x$ is a one-to-one function from $(0,1)$ to $(0,1]$. Finding a one-to-one function from $(0,1]$ to $(0,1)$ is also not difficult. The function $g(x) = x/2$ is clearly one-to-one and maps $(0,1]$ to $(0,1/2] \subset (0,1)$. As we have found one-to-one functions from $(0,1)$ to $(0,1]$ and from $(0,1]$ to $(0,1)$, the Schröder-Bernstein theorem tells us that $|(0,1)| = |(0,1]|$.■

Definition

We say that a function is **computable** if there is a computer program in some programming language that finds the values of this function. If a function is not computable we say it is **uncomputable**.

To show that there are uncomputable functions, we need to establish two results. First, we need to show that the set of all computer programs in any particular programming language is countable. This can be proved by noting that a computer programs in a particular language can be thought of as a string of characters from a finite alphabet (see **Exercise 44**). Next, we show that there are uncountably many different functions from a particular countably infinite set to itself. In particular, **Exercise 45** shows that the set of functions from the set of positive integers to itself is uncountable. This is a consequence of the uncountability of the real numbers between 0 and 1 (as shown in previous example). Putting these two results together (**Exercise 46**) shows that there are uncomputable functions.

Exercise Set (1.2)

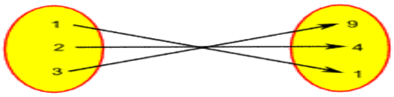
1- The function *f* is defined on the real numbers by $f(x) = 2 + x - x^2$. What is the value of $f(-3)$?

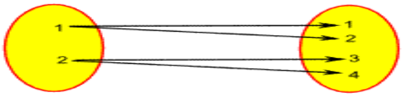
2- The function g is defined on the real numbers by $g(x) = (x^2 + 1)(3x - 5)$. What is the value of $g(4)$?

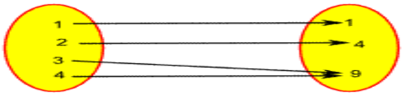
3- The function *f* is defined on the real numbers by $f(x) = x^2 - x - 10$. If $f(a) = -4$, what is the value of *a*?

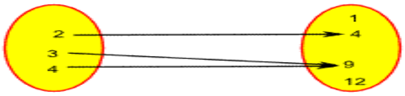
4-The function *f* is defined on the real numbers by $f(x) = 2x^2 - 5x + 12$. If $f(k) = 10$, what is the value of *k*?

5- Which one of the following relations is not a function?

A


B


C


D


6-What function is defined by the set of ordered pairs $\{\ldots,(-2,-5),(-1,-8),(0,-9),(1,-8),(2,-5),\ldots\}$?

Choose.

(a) $f(x) = x^2 - 9$ on the set of integers;

(b) $f(x) = x^2 - 9$ on the set of whole numbers;

(c) $f(x) = x^2 - 9$ on the set of real numbers;

(d) There is no such function.

7-Here is a set of ordered pairs:

$\{\ldots,(-2,7),(-1,1),(0,-1),(1,1),(2,7),\ldots\}$

Which function satisfies them?

(a) $f(x) = 2x - 1$ on the set of integers;

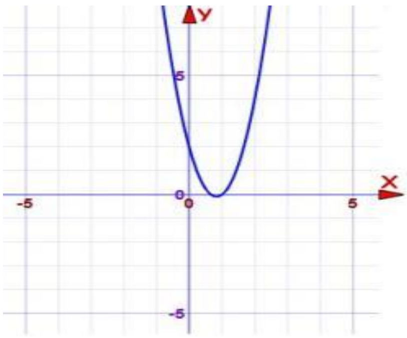
(b) $f(x) = -6x - 5$ on the set of integers;

(c) $f(x) = x^2 + 3$ on the set of integers;

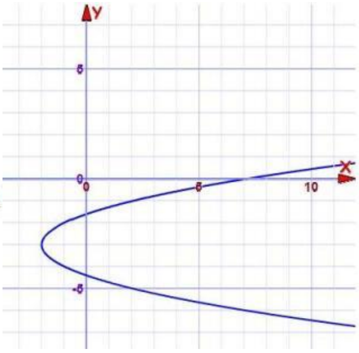
(d) $f(x) = 2x^2 - 1$ on the set of integers.

8-Which one of the following is not a function?

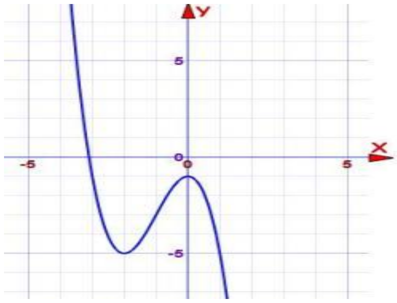
(a)



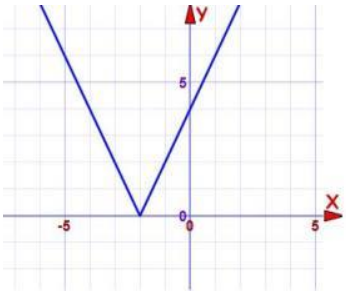
(b)



(c)

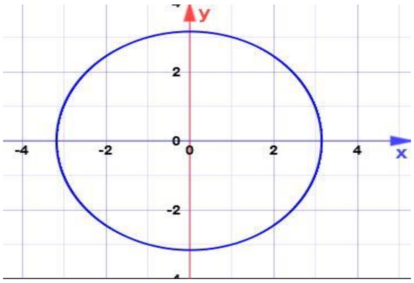


(d)

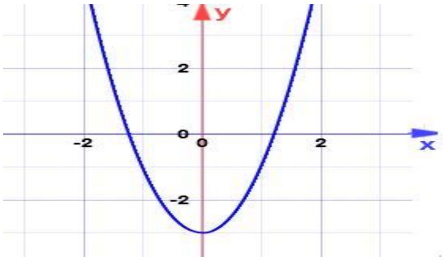


9-Which one of the following is not a function?

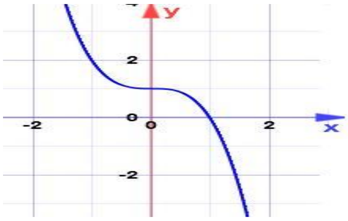
(a)



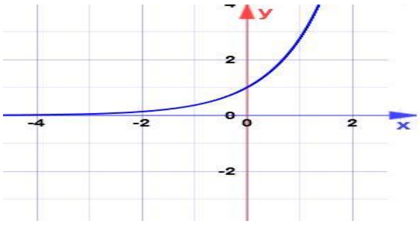
(b)



(c)



(d)



10- Why is f not a function from $\mathbb{R}$ to $\mathbb{R}$ if

(a) $f(x) = \frac{1}{x}$?; (b) $f(x) = \sqrt{x}$?; (c) $f(x) = \pm\sqrt{x^2 + 1}$?.

11- Determine whether f is a function from $\mathbb{Z}$ to $\mathbb{R}$ if

(a) $f(n) = \pm n$; (b) $f(n) = \sqrt{n^2 + 1}$; (c) $f(n) = \frac{1}{(n^2-4)}$.

12- Find the domain and range of these functions

(a) The function that assigns to each nonnegative integer its last digit;

(b) The function that assigns the next largest integer to a positive integer;

(c) The function that assigns to a bit string the number of one bits in the string.

(d) The function that assigns to a bit string the number of bits in the string;

(e) The function that assigns to each pair of positive integers the maximum of these two integers.

13- Find these values

(a) $\lfloor 1.1 \rfloor$; (b) $\lceil 1.1 \rceil$; (c) $\lfloor -0.1 \rfloor$; (d) $\lceil -0.1 \rceil$; (e) $\lceil 2.99 \rceil$;

(f) $\lceil -2.99 \rceil$; (g) $\left\lfloor \frac{1}{2} + \left\lceil \frac{1}{2} \right\rceil \right\rfloor$; (h) $\left\lceil \left\lfloor \frac{1}{2} \right\rfloor + \left\lceil \frac{1}{2} \right\rceil + \frac{1}{2} \right\rceil$.

14- Determine whether each of these functions from $\mathbb{Z}$ to $\mathbb{Z}$ is one-to- one.

(a) $f(n) = n - 1$; (b) $f(n) = n^2 + 1$;

(c) $f(n) = n^3$; (d) $f(n) = \left\lceil \frac{n}{2} \right\rceil$.

15- Determine whether the function $f: \mathbb{Z} \times \mathbb{Z} \rightarrow \mathbb{Z}$ is onto if

(a) $f(m, n) = m + n$; (b) $f(mn) = m^2 + n^2$;

(c) $f(m, n) = |n|$; (d) $f(m, n) = m$; (e) $f(m, n) = m - n$.

16- Determine whether each of these functions is a bijection from $\mathbb{R}$ to $\mathbb{R}$**:**

(a) $f(x) = 2x + 1$; (b) $f(x) = x^2 + 1$;

(c) $f(x) = x^3$; (d) $f(x) = \frac{(x^2+1)}{x^2+2}$;

17- Find $f \circ g$ and $g \circ f$, where $f(x) = x^2 + 1$ and $g(x) = x + 2$ are functions from $\mathbb{R}$ to $\mathbb{R}$**.**

18- Let $f(x) = ax + b$ and $g(x) = cx + d$ where a, b, c and d are constants.

Determine for which constants a, b, c and d it is true that

$$f \circ g = g \circ f.$$

19- Show that the function $f(x) = a x + b$ from $\mathbb{R}$ to $\mathbb{R}$ is invertible, where a and b are constants with $a \neq 0$, and find the inverse of f.

20- Let f be a function from the set A to the set B, let S and T be subsets of A. show that

(a)$f(S \cup T) = f(S) \cup f(T)$; (b)$f(S \cap T) \subseteq f(S) \cap f(T)$.

21- Let f be the function from $\mathbb{R}$ to $\mathbb{R}$ defined by $f(x) = x^2$. Find

(a) $f^{-1}(\{1\})$; (b) $f^{-1}(\{x: 0 < x < 1\})$; (c) $f^{-1}(\{x: x > 4\})$.

22- Suppose that f is a function from A to B, where A and B are finite sets with $|A| = |B|$. Show that f is one-to-one if and only if it is onto.

23- Determine whether each of these sets is countable or uncountable. For those that are countable, exhibit a one-to-one correspondence between the set of natural numbers and the set.

(a) the integers greater than 10.

(b) the odd negative integers.

(c) the real numbers between 0 and 2.

(d) integers that are multiple of 10.

(e) all positive rational numbers that cannot be written with denominators less than 4.

(f) all bit strings not containing the bit 0.

24- Write an equation to represent the function from the following table

x	y
-2	4
-1	1
0	0
1	1
2	4

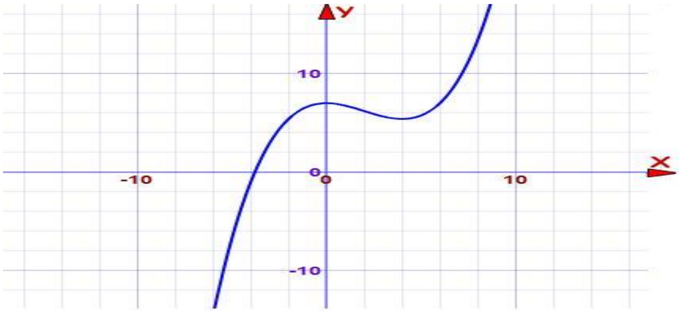
(a) $y = -2x$; (b) $y = 2x$; (c) $y = x + 2$; (d) $y = x^2$.

25 - Write an equation to represent the function from the following table of values:

x	y
-1	3
0	2
1	1
2	0
3	-1
4	-2

(a) $y = -x + 2$; (b) $y = x - 2$; (c) $y = x + 4$; (d) $y = x - 4$.

26- The following shows part of graph of the function $f(x) = 0.05x^3 - 0.3x^2 + 7$.



What are the Domain and Range of f?

(a) Domain = $\mathbb{R}$, Range = $\mathbb{R}$;

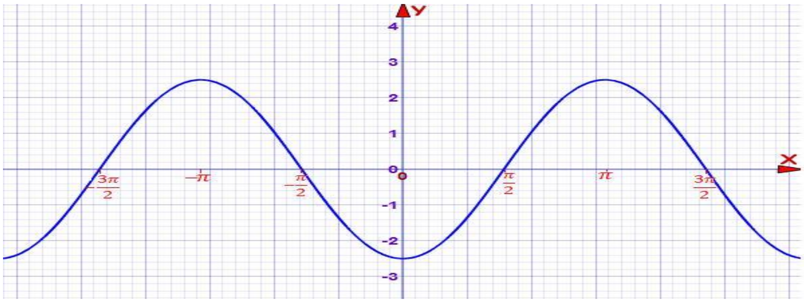
(b) Domain = $\{x \in \mathbb{R} \mid -10 < x < 10\}$, Range = $\mathbb{R}$;

(c) Domain = $\mathbb{R}$; Range = $\{y \in \mathbb{R} \mid -16 < y < 16\}$;

(d) Domain = $\{x \in \mathbb{R} \mid -10 < x < 10\}$,

Range = $\{y \in \mathbb{R} \mid -16 < y < 16\}$.

27- The following shows part of graph of the function $f(x) = 2.5\sin\left(x - \frac{\pi}{2}\right)$



What are the Domain and Range of *f*?

(a) Domain = $\mathbb{R}$, Range = $\mathbb{R}$;

(b) Domain = $\{x \in \mathbb{R} \mid -2\pi \le x \le 2\pi\}$,

Range = $\{y \in \mathbb{R} \mid -2.5 \le y \le 2.5\}$;

(b) Domain = $\mathbb{R}$, Range = $\{y \in \mathbb{R} \mid -1 \le y \le 1\}$;

(d) Domain = $\mathbb{R}$, Range = $\{y \in \mathbb{R} \mid -2.5 \le y \le 2.5\}$.

28- If $f(x) = \ln x$ and $g(x) = x + 1$, what is the domain of $(f \circ g)(x)$?

(a) $\{x \in \mathbb{R} \mid x \ge -1\}$; (b) $\{x \in \mathbb{R} \mid x > -1\}$;

(c) $\{x \in \mathbb{R} \mid x > 0\}$; (d) $\mathbb{R}$.

29- The function $f(x) = x^2$ is defined from $\mathbb{R}$ to $\mathbb{R}$. What is the Codomain?

(a) $\{y \in \mathbb{R} | y \geq 0\}$; (b) $\{y \in \mathbb{R} | y > 0\}$; (c) $\{y \in \mathbb{R} | y \neq 0\}$; (d) $\mathbb{R}$.

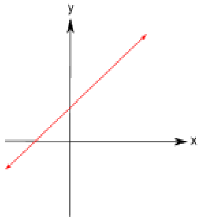
30- The function $f(x) = e^x + 3$ is defined from $\mathbb{R}$ to $\mathbb{R}$. What is the Range?

(a) $\{y \in \mathbb{R} | y \geq 3\}$; (b) $\{y \in \mathbb{R} | y > 3\}$; (c) $\{y \in \mathbb{R} | y \geq 0\}$; (d) $\mathbb{R}$.

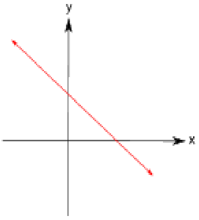
31- The function $f(x) = \text{floor}(x)$ is defined from $\mathbb{R}$ to $\mathbb{R}$. What is the Range?

32- Which one of these graphs does not illustrate a function?

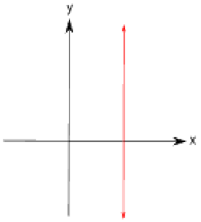
A



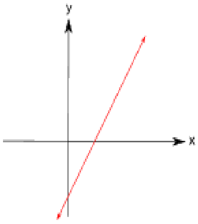
B



C



D



33- The following sets of ordered pairs represent relations from the set X to the set Y. Which one is not a function?

(a) $\{(1,2),(2,4),(3,6),(4,8)\}$;

(b) $\{(1,2),(1,4),(1,6),(1,8)\}$;

(c) $\{(1,1),(2,4),(3,9),(4,16)\}$;

(d) $\{(1,1),(2,3),(3,5),(4,7)\}$.

34- If $f(x) = \sqrt{x-3}$ and $g(x) = \sqrt{4-x}$, what is the domain of the function $(f+g)(x)$?

35- If $f(x) = x + 1$ and $g(x) = \sqrt{1 - x}$, what is the domain of the function $(f/g)(x)$?

36- If $f(x) = \frac{1}{x-2}$ and $g(x) = \frac{1}{x+2}$, what is the domain of the function $(f - g)(x)$?

37- If $f(x) = 3x - 15$ and $g(x) = \sqrt{x - 5}$, then what is the function $(f/g)(x)$ and what is its domain?

38- If $f(x) = x^2 + x - 6$ and $g(x) = \frac{1}{x+3}$, then what is the function $(f \cdot g)(x)$ and what is its domain?

39- If $f(x) = x^2 + 3$ and $g(x) = \sqrt{x - 3}$, then what is the function $(f \circ g)(x)$ and what is its domain?

40. f and g are both defined on the set of real numbers, $f(x) = x^2$ and $g(x) = x + 2$. For what value of x does $(f \circ g)(x) = (g \circ f)(x)$?

41 . f and g are both defined on the set of real numbers and c is a constant, where $f(x) = cx - 3, g(x) = cx + 5$.

If $(f \circ g)(x) = (g \circ f)(x)$ for all x, what is the value of c?

42 . If $f(x) = x + 2$ and $g(x) = \frac{1}{x-2}$, then what is the function $(g \circ f)(x)$ and what is its domain?

43 . $f(x) = x^3$ and $g(x) = \frac{1}{x} + 1$. The domain for $f = \mathbb{R}$ and the domain for $g = \{x \in \mathbb{R} \mid x \neq 0\}$. For what value of x does $(f \circ g)(x) = (g \circ f)(x)$?

44. Show that the set of all computer programs in a particular programming language is countable.

[*Hint:* A computer program written in a programming language can be thought of as a string of symbols from a finite alphabet.]

45. Show that the set of functions from the positive integers to the set $\{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}$ is uncountable.

[*Hint:* First set up a one-to-one correspondence between the set of real numbers between 0 and 1 and a subset of these functions. Do this by associating to the real number $0. d_1 d_2 \dots d_n \dots$ the function f with $f(n) = d_n$.]

46. We say that a function is **computable** if there is a computer program that finds the values of this function. Use Exercises 44 and 45 to show that there are functions that are not computable.

1.3 Relations

Definition.

A subset R of the Cartesian product $A \times B$ is called a **relation** from the set A to the set B. The elements of R are ordered pairs, where the first element belongs to A and the second element to B. We use the notation aRb to denote that $(a, b) \in R$ which means a is said to be related to b by R. Moreover, when $(a, b) \notin R$ we mean, a is not related to b by R.

Example

$R = \{(a, 0), (a, 1), (a, 3), (b, 1), (b, 2)\}$ is a relation from the set $\{a, b, c\}$ to the set $\{0, 1, 2, 3\}$. ■

The Cartesian product, $A \times B$ and $B \times A$ are not equal, unless $A = \phi$ or $B = \phi$ (so that $A \times B = \phi$) or unless $A = B$.

Definition

In mathematics, a **binary relation** on a set A is a collection of ordered pairs of elements of A. In other words, it is a subset of the Cartesian product $A^2 = A \times A$. More generally, a binary relation between two sets A and B is a subset of $A \times B$.

Example

The "divides" relation between the set of prime numbers $\mathbb{P}$ and the set of integers $\mathbb{Z}$, in which every prime p is associated with every integer z that is a multiple of p (and not with any integer that is not a multiple of p). In this relation, for instance, the prime

2 is associated with numbers that include $-4, 0, 6, 10$, but not 1 or 9; and the prime 3 is associated with numbers that include 0, 6, and 9, but not 4 or 13. ■

Binary relations are used in many branches of mathematics to model concepts like:

"is greater than", "is equal to", and "divides" in arithmetic,

"is congruent to" in geometry,

"is adjacent to" in graph theory,

"is orthogonal to" in linear algebra.

The concept of function is defined as a special kind of binary relation.

A binary relation is the special case $n = 2$ of an *n*-ary relation $R \subseteq A_1 \times \ldots \times A_n$, that is, a set of *n*-tuples where the *j*th component of each *n*-tuple is taken from the *j*th domain A_j of the relation.

Example

Some Examples of Relations include:

$\{ (0,1)\, , (55,22), (3, -50) \}$;

$\{(0, 1)\, , (5, 2), (-3, 9) \}$;

$\{ (-1,7)\, , (1, 7), (33, 7), (32, 7) \}$.■

Definition

The domain of the relation is the set of all the *first* numbers of the ordered pairs. In other words, the domain is all of the *x*-

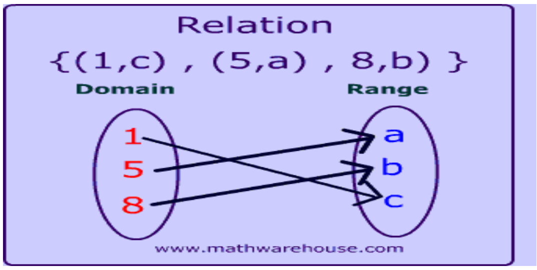
values. **The range** is the set of the *second* numbers in each pair, or the y-values.

Example

The relation $\{(0,1),(3,22),(90,34)\}$ its domain is $\{\,0,3,\ 90\,\}$ and the range is $\{\,1,\ 22,\ 34\,\}$. ■

Relations are often represented using arrow charts connecting the domain and range elements.

Example



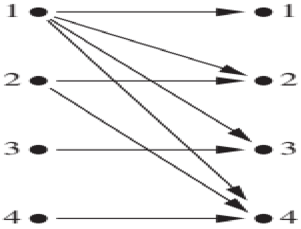
Example

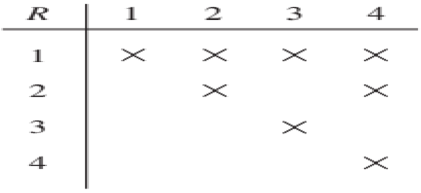
Let A be the set $\{1,2,3,4\}$. Which ordered pairs are in the relation $R = \{(a,b) \mid a \text{ divides } b\}$?

Solution

Because (a,b) is in R if and only if a and b are positive integers not exceeding 4 such that a divides b, we see that

$$R = \{(1,1),(1,2),(1,3),(1,4),(2,2),(2,4),(3,3),(4,4)\}.$$





R	1	2	3	4
1	×	×	×	×
2		×		×
3			×	
4				×

The pairs in this relation are displayed both graphically and in tabular form in the above figure. ■

The following are relations on an infinite set.

Example

Consider these relations on the set of integers:

$R_1 = \{(a, b) \mid a \leq b\}$,

$R_2 = \{(a, b) \mid a > b\}$,

$R_3 = \{(a, b) \mid a = b \, or \, a = -b\}$,

$R_4 = \{(a, b) \mid a = b\}$,

$R_5 = \{(a, b) \mid a = b + 1\}$,

$R_6 = \{(a, b) \mid a + b \leq 3\}$.

Which of these relations contain each of the pairs $(1, 1), (1, 2), (2, 1), (1, -1)$, and $(2, 2)$?

Solution

The pair $(1, 1)$ is in R_1, R_3, R_4, and R_6;

The pair $(1, 2)$ is in R_1 and R_6;

The pair $(2, 1)$ is in R_2, R_5, and R_6;

The pair $(1, -1)$ is in R_2, R_3, and R_6;

The pair $(2, 2)$ is in R_1, R_3, and R_4. ■

It is not hard to determine the number of relations on a finite set, because a relation on a set A is simply a subset of $A \times A$.

Example

How many relations are there on a set with *n* elements?

Solution

A relation on a set *A* is a subset of $A \times A$. Because $A \times A$ has n^2 elements when *A* has *n* elements, and a set with *m* elements has 2^m subsets, there are 2^{n^2} subsets of $A \times A$. Thus, there are 2^{n^2} relations on a set with *n* elements. ■

Example

There are $2^{3^2} = 2^9 = 512$ relations on the set $\{a, b, c\}$.■

●Functions as Relations

Recall that a function *f* from a set *A* to a set *B* assigns exactly one element of *B* to each element of *A*. The graph of *f* is the set of ordered pairs (a, b) such that $b = f(a)$. Because the graph of *f* is a subset of $A \times B$, it is a relation from *A* to *B*. Moreover, the graph of a function has the property that every element of *A* is the first element of exactly one ordered pair of the graph.

Conversely, if *R* is a relation from *A* to *B* such that every element in *A* is the first element of exactly one ordered pair of *R*, then a function can be defined with *R* as its graph. This can be done by assigning to an element *a* of *A* the unique element $b \in B$ such that $(a, b) \in R$. A relation can be used to express a one-to-many relationship between the elements of the sets A and B, where an element of A may be related to more than one element of B. A function represents a relation where exactly one element of B is related to each element of A. Relations are a generalization of graphs of functions; they can be used to express a much wider class of relationships between sets. (Recall that the graph of the function f from A to B is the set of ordered pairs $(a, f\ (a))\ for\ a\ \in\ A$.)

***What makes a relation a* function?**

As soon as an element in the domain repeats, the relation is not a function.

Example.

Which relations below are functions?

Relation #1 $\{\ (-1,2), (-4,51), (1,2), (8, -51)\ \}$;

Relation #2 $\{(13, 14), (13, 5), (16,7), (18,13)\ \}$;

Relation #3 $\{\ (3,90), (4,54), (6,71), (8,90)\ \}$.

Solution

Both Relation #1 and Relation #3 are functions, but Relation #2 is not a functions as the x-place 13 appeared twice. ■

Practice

For the following relation to be a function, X cann**ot** be what values?

$$\{(8\ , 11), (34,5), (6,17), (X\ ,22)\ \}$$

●Properties of Relations

In some relations an element is always related to itself. For instance, let R be the relation on the set of all people consisting of pairs (x, y), where x and y have the same mother and the same father. Then xRx for every person x, which is defined as follows:

Definition

A relation R on a set A is called *reflexive* if $(a, a) \in R$ for every element $a \in A$.◀

Example

Let $A = \{a, b, c\}$ and $R = \{(a, a), (b, b), (c, c)\}$. Then R is a reflexive relation in A. ■

Example

'Equality' is a reflexive relation, since an element equals itself. ■

In some relations an element is related to a second element if and only if the second element is also related to the first element. The relation consisting of pairs (x, y), where x and y are students at your school with at least one common class has this property, which is defined as follows:

Definition

A relation R on a set A is called *symmetric* if $(b, a) \in R$ whenever $(a, b) \in R$, for all $a, b \in A$.◀

Example

Let R be relation ‘is perpendicular to’ in the set of all straight lines, then R is a symmetric relation. ■

Other relations have the property that if an element is related to a second element, then this second element is not related to the first. The relation consisting of the pairs (x, y), where x and y are students at your school, where x has a higher grade point average than y has this property, which is defined as follows:

Definition

A relation R on a set A such that for all $a, b \in A$, if $(a, b) \in R$ and $(b, a) \in R$, then $a = b$ is called *antisymmetric*.◀

Example.

Let ℕ be the set of Natural Numbers R be a relation in ℕ, defined by ‘a is a divisor’ of b, i.e., aRb if a divides b then R is antisymmetric since a divides b and b divides $a \Rightarrow a = b$. ■

Let R be the relation consisting of all pairs (x, y) of students at your school, where x has taken more credits than y. Suppose that x is related to y and y is related to z. This means that x has taken more credits than y and y has taken more credits than z. We can conclude that x has taken more credits than z, so that x is related to z. What we have shown is that R has the transitive property, which is defined as follows:

Definition.

A relation R on a set A is called *transitive* if whenever $(a, b) \in R$ and $(b, c) \in R$, then $(a, c) \in R$, for all $a, b, c \in A$.◀

Example.

Let A be the set of straight lines in a plane and R be a relation in A defined by ‘is parallel to’. Then R is a transitive relation in A. ■

Example.

Let $A = \{1, 2, 3\}$ and $R = \{(1, 1), (2, 2), (2, 3), (3, 2), (3, 3)\}$ then R is transitive. ■

Example.

Consider these relations on the set of integers:

$R_1 = \{(a, b) \mid a \leq b\};$

$R_2 = \{(a, b) \mid a > b\};$

$R_3 = \{(a, b) \mid a = b \, or \, a = -b\};$

$R_4 = \{(a, b) \mid a = b\};$

$R_5 = \{(a, b) \mid a = b + 1\};$

$R_6 = \{(a, b) \mid a + b \leq 3\}.$

Which of these relations are:

(a) reflexive; (b) symmetric; (c) antisymmetric; (d) transitive.?

Solution.

(a) The reflexive relations are R_1 (because $a \leq a$ for every integer a), R_3, and R_4. For each of the other relations in this example it is easy to find a pair of the form (a, a) that is not in the relation.

(b) The relations R_3, R_4, and R_6 are symmetric. R_3 is symmetric, for if $a = b$ or $a = -b$, then $b = a$ or $b = -a$. R_4 is symmetric because $a = b$ implies that $b = a$. $R6$ is symmetric because $a + b \leq 3$ implies that $b + a \leq 3$. The reader should verify that none of the other relations is symmetric.

(c) The relations R_1, R_2, R_4, and R_5 are antisymmetric. R_1 is antisymmetric because the inequalities $a \leq b$ and $b \leq a$ imply that $a = b$. R_2 is antisymmetric because it is impossible that $a > b$ and $b > a$. R_4 is antisymmetric, because two elements are related with respect to R_4 if and only if they are equal. R_5 is antisymmetric because it is impossible that $a = b + 1$ and $b = a + 1$. The reader should verify that none of the other relations is antisymmetric.

(d) The relations R_1, R_2, R_3, and R_4 are transitive. R_1 is transitive because $a \leq b$ and $b \leq c$ imply that $a \leq c$. R_2 is transitive because $a > b$ and $b > c$ imply that $a > c$. R_3 is transitive because $a = \pm b$ and $b = \pm c$ imply that $a = \pm c$. R_4 is clearly transitive, as the reader should verify. R_5 is not transitive because $(2, 1)$ and $(1, 0)$ belong to R_5, but $(2, 0)$ does not. R_6 is not transitive because $(2, 1)$ and $(1, 2)$ belong to R_6, but $(2, 2)$ does not. ■

Example.

Consider the following relations on $\{1, 2, 3, 4\}$:

$R_1 = \{(1, 1), (1, 2), (2, 1), (2, 2), (3, 4), (4, 1), (4, 4)\}$,

$R_2 = \{(1, 1), (1, 2), (2, 1)\}$,

$R_3 = \{(1, 1), (1, 2), (1, 4), (2, 1), (2, 2), (3, 3), (4, 1), (4, 4)\}$,

$R_4 = \{(2, 1), (3, 1), (3, 2), (4, 1), (4, 2), (4, 3)\}$,

R_5

$= \{(1, 1), (1, 2), (1, 3), (1, 4), (2, 2), (2, 3), (2, 4), (3, 3), (3, 4), (4, 4)\}$

$R_6 = \{(3, 4)\}$.

Which of these relations are:

(a) reflexive; (b) symmetric; (c) antisymmetric; (d) transitive.?

Solution.

(a) The relations R_2 and R_5 are reflexive because they both contain all pairs of the form (a, a), namely, $(1, 1), (2, 2), (3, 3)$, and $(4, 4)$. The other relations are not reflexive because they do not contain all of these ordered pairs. In particular, R_1, R_2, R_4, and R_6 are not reflexive because $(3, 3)$ is not in any of these relations.

(b) The relations R_2 and R_3 are symmetric, because in each case (b, a) belongs to the relation whenever (a, b) does. For R_2, the only thing to check is that both $(2, 1)$ and $(1, 2)$ are in the relation. For R_3, it is necessary to check that both $(1, 2)$ and $(2, 1)$ belong to the relation, and $(1, 4)$ and $(4, 1)$ belong to the relation. The reader should verify that none of the other relations is symmetric. This is done by finding a pair (a, b) such that it is in the relation but (b, a) is not.

(c) R_4, R_5, and R_6 are all antisymmetric. For each of these relations there is no pair of elements a and b with $a \neq b$ such that both (a, b) and (b, a) belong to the relation. The reader should verify that none of the other relations is antisymmetric. This is done by finding a pair (a, b) with $a \neq b$ such that (a, b) and (b, a) are both in the relation.

(d) R_4, R_5, and R_6 are transitive. For each of these relations, we can show that it is transitive by verifying that if (a, b) and (b, c) belong to this relation, then (a, c) also does. For instance, R_4 is transitive, because $(3, 2)$ and $(2, 1)$, $(4, 2)$ and $(2, 1)$, $(4, 3)$ and $(3, 1)$, and $(4, 3)$ and $(3, 2)$ are the only such sets of pairs, and $(3, 1)$, $(4, 1)$, and $(4, 2)$ belong to R_4. The reader should verify that R_5 and R_6 are transitive. R_1 is not transitive because $(3, 4)$ and $(4, 1)$ belong to R_1, but $(3, 1)$ does not. R_2 is not transitive because $(2, 1)$ and $(1, 2)$ belong to R_2, but $(2, 2)$ does not. R_3 is not transitive because $(4, 1)$ and $(1, 2)$ belong to R_3, but $(4, 2)$ does not.■

Example.

Is the “divides” relation on the set of positive integers:

(a) reflexive; (b) symmetric; (c) antisymmetric; (d) transitive.?

Solution.

(a) Because $a|a$ whenever a is a positive integer, the “divides” relation is reflexive. (Note that if we replace the set of positive integers with the set of all integers the relation is not reflexive because by definition 0 does not divide 0.)

(b) This relation is not symmetric because $1|2$, but $2 \nmid 1$.

(c) It is antisymmetric, for if a and b are positive integers with $a|b$ and $b|a$, then $a = b$ (the verification of this is left as an exercise for the reader).

(d) Suppose that a divides b and b divides c. Then there are positive integers k and l such that $b = ak$ and $c = bl$. Hence, $c = a(kl)$, so a divides c. It follows that this relation is transitive. ■

Example.

How many reflexive relations are there on a set with *n* elements?

Solution.

A relation R on a set A is a subset of $A \times A$. Consequently, a relation is determined by specifying whether each of the n^2 ordered pairs in $A \times A$ is in R. However, if R is reflexive, each of the *n* ordered pairs (a, a) for $a \in A$ must be in *R*. Each of the other $n(n - 1)$ ordered pairs of the form (a, b), where $a \neq b$, may or may not be in *R*. Hence, by the product rule for counting, there are $2^{n\,(n-1)}$ reflexive relations [this is the number of ways to choose whether each element (a, b), with $a \neq b$, belongs to R]. ■

● Operations on binary relations

Because relations from A to B are subsets of $A \times B$, two relations from A to B can be combined in any way two sets can be combined. Consider the following examples.

Example.

Let $A = \{1, 2, 3\}$ and $B = \{1, 2, 3, 4\}$. The relations

$R_1 = \{(1, 1), (2, 2), (3, 3)\}$ and

$R_2 = \{(1, 1), (1, 2), (1, 3), (1, 4)\}$

can be combined to obtain

$R_1 \cup R_2 = \{(1, 1), (1, 2), (1, 3), (1, 4), (2, 2), (3, 3)\}$,

$R_1 \cap R_2 = \{(1, 1)\}$,

$R_1 - R_2 = \{(2, 2), (3, 3)\}$,

$R_2 - R_1 = \{(1, 2), (1, 3), (1, 4)\}$. ■

Example.

Let A and B be the set of all students and the set of all courses at a school, respectively. Suppose that R_1 consists of all ordered pairs (a, b), where a is a student who has taken course b, and R_2 consists of all ordered pairs (a, b), where a is a student who requires course b to graduate.

What are the relations $R_1 \cup R_2$, $R_1 \cap R_2$, $R_1 \oplus R_2$, $R_1 - R_2$, and $R_2 - R_1$?

Solution.

The relation $R_1 \cup R_2$ consists of all ordered pairs (a, b), where a is a student who either has taken course b or needs course b to graduate.

$R_1 \cap R_2$ is the set of all ordered pairs (a, b), where a is a student who has taken course b and needs this course to graduate.

Also, $R_1 \oplus R_2$ consists of all ordered pairs (a, b), where student a has taken course b but does not need it to graduate or needs course b to graduate but has not taken it.

$R_1 - R_2$ is the set of ordered pairs (a, b), where a has taken course b but does not need it to graduate; that is, b is an elective course that a has taken. $R_2 - R_1$ is the set of all ordered pairs (a, b), where b is a course that a needs to graduate but has not taken. ■

Example.

Let R_1 be the "less than" relation on the set of real numbers and let R_2 be the "greater than" relation on the set of real numbers, that is, $R_1 = \{(x, y) \mid x < y\}$ and $R_2 = \{(x, y) \mid x > y\}$.

What are the relations

$R_1 \cup R_2, R_1 \cap R_2, , R_1 - R_2, R_2 - R_1$, and $R_1 \oplus R_2$?

Solution.

We note that $(x, y) \in R_1 \cup R_2$ if and only if $(x, y) \in R_1$ or $(x, y) \in R_2$. Hence, $(x, y) \in R_1 \cup R_2$ if and only if $x < y$ or

$x > y$. Because the condition $x < y$ or $x > y$ is the same as the condition $x \neq y$, it follows that $R_1 \cup R_2 = \{(x, y) \mid x \neq y\}$. In other words, the union of the "less than" relation and the "greater than" relation is the "not equals" relation. Next, note that it is impossible for a pair (x, y) to belong to both R_1 and R_2 because it is impossible that $x < y$ and $x > y$. It follows that $R_1 \cap R_2 = \phi$.

We also see that $R_1 - R_2 = R_1, R_2 - R_1 = R_2$,

and $R_1 \oplus R_2 = R_1 \cup R_2 - R_1 \cap R_2 = \{(x, y) \mid x \neq y\}$.■

Definition.

Let *R* be a relation from a set *A* to a set *B* and *S* a relation from *B* to a set *C*. The *composite* of *R* and *S* is the relation consisting of ordered pairs (a, c), where $a \in A$, $c \in C$, and for which there exists an element $b \in B$ such that $(a, b) \in R$ and $(b, c) \in S$. We denote the composite of *R* and *S* by $S \circ R$.◀

Computing the composite of two relations requires that we find elements that are the second element of ordered pairs in the first relation and the first element of ordered pairs in the second relation, as the following examples.

Example.

What is the composite of the relations *R* and *S*, where *R* is the relation from $\{1, 2, 3\}$ to $\{1, 2, 3, 4\}$ with $R = \{(1, 1), (1, 4),$

$(2,3), (3,1), (3,4)\}$ and S is the relation from $\{1,2,3,4\}$ to $\{0,1,2\}$ with $S = \{(1,0),(2,0),(3,1),(3,2),(4,1)\}$?

Solution.

$S \circ R$ is constructed using all ordered pairs in R and ordered pairs in S, where the second element of the ordered pair in R agrees with the first element of the ordered pair in S. For example, the ordered pairs $(2,3)$ in R and $(3,1)$ in S produce the ordered pair $(2,1)$ in $S \circ R$. Computing all the ordered pairs in the composite, we find $S \circ R = \{(1,0),(1,1),(2,1),(2,2),(3,0),(3,1)\}$. ■

Example.

Let R be the relation on the set of all people such that (a,b) in R if a is a parent of b. Then (a,c) in $R \circ R$, if (b,c) in R. This means that a is a grandparent of c.

● R^{n}

Definition.

The power R^n, for $n = 1,2,3,\ldots$ are defined by $R^1 = R$, and $R^{n+1} = R^n \circ R$ where R is a relation on the set A. ◄

The definition shows that $R^2 = R \circ R$, $R^3 = R^2 \circ R = (R \circ R) \circ R$, and so on.

Example

Let $A = \{2,4,6\}$ and $B = \{3,6,9\}$

$$A \times B$$
$$= \{(2,3),(2,6),(2,9),(4,3),(4,6),(4,9),(6,3),(6,6),(6,9)\}$$

Let R be a relation from A to B such that

$R^1 = \{ (2,3),(2,6),(4,3),(4,9),(6,6),(6,9) \}$.

$R^2 = \{ (2,6),(2,9),(6,6),(6,9) \}$

$R^3 = \{ (2,6),(2,9),(6,6),(6,9) \}$ etc.

Example

Let $R = \{(1,1),(2,1),(3,2),(4,3)\}$.

Find the powers $R^n, n = 2,3,4,\dots.$

Solution

Because $R^2 = R \circ R$, we find $R^2 = \{(1,1),(2,1),(3,1),(4,2)\}$. Furthermore, $R^3 = (R \circ R) \circ R = \{(1,1),(2,1),(3,1),(4,1)\}$. Additional computation shows that R^4 is the same as R^3, so $R^4 = \{(1,1),(2,1),(3,1),(4,1)\}$. It also follows that $R^n = R^3$ for $n = 5,6,7,\dots.$ ■

Definition.

A relation R on sets X and Y is said to be **contained** in a relation S on X and Y if R is a subset of S, that is, if $x\,R\,y$ always implies xSy.

In this case, if R and S disagree, R is said to be **smaller** than S. ◄

Example.

$>$ is contained in $\geq$.■

Theorem.

The relation R on a set A is transitive if and only if $R^n \subseteq R$ for $n = 1,2,3,\dots$ ▲

If R is a binary relation over A and B, then the following is a binary relation over B and A:

Definition

Let R be a relation from A to B. Then the relation $R^{-1} = \{ (b,a) \mid (a,b) \in R \}$ from B to A is called the inverse of R. ◀

Example

"is less than" (<) is the inverse of "is greater than" (>).■

Example

Let $A = \{1,2,3\}$, $B = \{4,5\}$ and $R = \{(1,4),(2,5),(3,5)\}$ be a relation from A to B. then $R^{-1} = \{(4,1),(5,2),(5,3)\}$. ■

Theorem.

A binary relation over a set is equal to its inverse if and only if it is symmetric. ▲

Definition

If R is a binary relation over X and Y, then the following too:

The **complement** R^c is defined as $x\ R^c\ y$ if not $x\ R\ y$.

Example

On real numbers, $\leq$ is the complement of $>$.■

Example

Let $A = \{1,2,3\}$ and $R = \{(1,1),(1,2),(1,3),(2,2),(3,3)\}$

Then $R^c = \{(2,1),(2,3),(3,1),(3,2)\}$. ■

The complement of the inverse is the inverse of the complement.

If a relation is symmetric, the complement is too.

● Equivalence Relations

Example

Let R be the relation on the set of real numbers such that aRb if and only if $a - b$ is an integer. Is R an equivalence relation?

Solution

Because $a - a = 0$ is an integer for all real numbers a, aRa for all real numbers a. Hence, R is reflexive. Now suppose that aRb. Then $a - b$ is an integer. Then $b - a = -(b - a)$ is an integer. Hence bRa. It follows that R is symmetric. If aRb and bRc, then $a - b$ and $b - c$ are integers. Therefore $a - c = (a - b) + (b - c)$ is also an integer. Hence aRc. Thus, R is transitive. Consequently, R is an equivalence relation. ■

Example (Congruence Modulo m).

Let m be a positive integer with $m > 1$. Show that the relation $R = \{(a, b): a \equiv b(mod\, m)\}$ is an equivalence relation on the set of integers. Where $a \equiv b(mod\, m)$ if and only if m divides $a - b$.

Solution

Note that $a - a = 0$ is divisible by m, because $0 = 0 \cdot m$. Hence $a \equiv a(mod\, m)$, so that congruence modulo m is reflexive. Now, suppose that $a \equiv b(mod\, m)$. Then $a - b$ is divisible by m, so $a - b = k\, m$, k is an integer. It follows that $b - a = (-k)\, m$, so that $b \equiv a(mod\, m)$. So congruence modulo m is symmetric. Suppose $a \equiv b(mod\, m)$ and $b = c\ (mod\, m)$. Then there are integers k and ℓ with $a - b = k\,m$ and $b - c = \ell m$. Adding these two equations shows that

$a - c = (a - b) + (b - c) = km + \ell m = (k + \ell)m$. Thus, $a \equiv c(mod\, m)$. Therefore, congruence modulo m is transitive. It follows that congruence modulo m is an equivalence relation. ■

Discussion.

If $a \equiv b(mod\, m)$, then by definition of congruence, $m|(a - b)$. This means that there is an integer k such that $a - b = km$, so that $a = b + km$. Conversely, if there is an integer k such that $a = b + km$, then $km = a - b$. Hence m divides $a - b$, so that $a \equiv b(mod\, m)$. From this discussion we can state that "Let m be a positive integer. The integers a and b are congruent modulo m if and only if there is an integer k such that $a = b + km$.

Example

Suppose ~ is relation on $\mathbb{N} \times \mathbb{N}$

$$(m,n)\sim(p,q) \Leftrightarrow m + q = p + n$$

Prove that ~ is an equivalence relation on $\mathbb{N} \times \mathbb{N}$.

Solution

Since $m + n = n + m$, then $(m,n)\sim(m,n)\forall m,n \in N$.

Therefore ~ is reflexive.

Let $(m,n)\sim(p,q)$. Then $m + q = p + n$ or $p + n = m + q$

Therefore $(p,q)\sim(m,n)$ and ~ is symmetric.

$(m,n)\sim(p,q)$ and $(p,q)\sim(r,s) \Longrightarrow$

$m+q=p+n$ and $p+s=q+r$.

Then $m+s=n+r \Longrightarrow (m,n)\sim(r,s)$ and $\sim$ is transitive.

Therefore, $\sim$ is an equivalence relation. ■

Example.

Let S be a relation on $\mathbb{R}$ defined as

$$x^2-y^2=2(y-x) \Leftrightarrow (x,y)\in S$$

Prove that S is an equivalence relation.

Solution

Not that $x^2-y^2=2(y-x) \Leftrightarrow x^2+2x=y^2+2y$.

Since $x^2+2x=x^2+2x \quad \forall x\in\mathbb{R} \Longrightarrow (x,x)\in S \Longrightarrow S$ is reflexive

Let $(x,y)\in S \Longrightarrow x^2+2x=y^2+2y \Longrightarrow$

$y^2+2y=\ x^2+2x \Longrightarrow (y,x)\ \in S$ is symmetric.

Since $(x,y),(y,z)\in S \Longrightarrow x^2+2x=\ y^2+2y$ and

$y^2+2y=\ z^2+2z \Longrightarrow x^2+2x=\ z^2+2z \Longrightarrow (x,z)\in S$

Hence S is transitive.

Therefore, S is an equivalence relation. ■

Example

Let $\sim$ be a relation on $\mathbb{Q}^+$ such that: $x\sim y \Leftrightarrow \frac{x}{y}\in\mathbb{Q}^+$

Prove that $\sim$ is an equivalence relation.

Solution

Since $\frac{x}{x}=1\in\mathbb{Q}^+ \Longrightarrow (x,x)\in\sim$ or $\sim$ is reflexive.

If $\frac{x}{y} \in \mathbb{Q}^+ \Longrightarrow \frac{y}{x} \in \mathbb{Q}^+ \Longrightarrow \sim$ is symmetric.

Since $\frac{x}{y}, \frac{y}{z} \in \mathbb{Q}^+ \Longrightarrow \frac{x}{z} = \frac{x}{y} . \frac{y}{z} \in \mathbb{Q}^+ \Longrightarrow \sim$ is transitive.

Hence $\sim$ is an equivalence relation. ■

● Equivalence Classes

Let *A* be the set of all students in your school who graduated from high school. Consider the relation *R* on *A* that consists of all pairs (x, y), where *x* and *y* graduated from the same high school. Given a student *x*, we can form the set of all students equivalent to *x* with respect to *R*. This set consists of all students who graduated from the same high school as *x* did. This subset of *A* is called an equivalence class of the relation.

Definition

Let *R* be an equivalence relation on a set *A*. The set of all elements that are related to an element *a* of *A* is called **equivalence class** of *a*. The equivalence class of *a* with respect to R is denoted by $[a]_R$. When only one relation is under consideration, we can delete the subscript *R* and write $[a]$ for this equivalence class.

In other words, if *R* is an equivalence relation on a set *A*, the equivalence class of the element *a* is $[a]_R = \{s : (a, s) \in R\}$. If $b \in [a]_R$, then *b* is called a **representative** of this equivalence class. Any element of a class can be used as a representative of this class. ◀

Example

What are the equivalence classes of 0 and 1 for congruence modulo 4?

Solution

The equivalence class of 0 contains all integers *x* such that

$$x \equiv 0 \pmod{4},$$

i.e. the integers divisible by 4. So, the equivalence class of 0 is

$$[\,0\,]_4 \;=\; \{\ldots,-12,-8,-4,0,4,8,12,\ldots\}.$$

The equivalence class of 1 contains all integers y such that

$$y \;\equiv\; 1\,(\bmod 4\,),$$

i.e. the integers with remainder 1 when divided by 4. Hence, the equivalence class of 1 is

$$[\,1\,]4 \;=\; \{\ldots,-11,-7,-3,1,5,9,13,\ldots\}.$$■

Definition

The equivalence classes of the relation congruence modulo m are called the **congruent classes modulo** m. The congruence class of an integer *a* modulo m is denoted by $[a]_m$. Hence,

$$[a]_m = \{\ldots, a-2m, a-m, a, a+m, a+2m, \ldots\}.$$◄

Example

From the above example it follows that

$[0]_4 = \{\ldots,\ -8,\ -4,\ 0,\ 4,\ 8,\ \ldots\};$

$[1]_4 = \{\ldots,-7,-3,\ 1,\ 5,\ 9,\ \ldots\};$

$[2]_4 = \{\ldots,\ -6,\ -2,\ 2,\ 6,\ 10,\ \ldots\}.$■

Example

We have proved that $\sim$ is an equivalence relation on $\mathbb{N} \times \mathbb{N}$ defined by $(m,n)\sim(p,q) \Leftrightarrow m+q=p+n$. So, we can find $[(1,1)]$ and $[(3,4)]$ as follows.

$$[(1,1)] = \{(a,b) \in \mathbb{N} \times \mathbb{N}: (a,b)\sim(1,1)\}$$
$$= \{(a,b) \in \mathbb{N} \times \mathbb{N}: a+1=b+1\}$$
$$= \{(a,a): a \in \mathbb{N}\} = \{(0,0),(1,1),(2,2),\ldots.\}$$
$$[(3,4)] = \{(a,b) \in \mathbb{N} \times \mathbb{N}: a+4=b+3\}$$
$$= \{(a,b) \in \mathbb{N} \times \mathbb{N}: b=a+1\}$$
$$= \{(0,1),(1,2),(2,3),\ldots..\}. \blacksquare$$

Example

Let S be the equivalence relation on $\mathbb{R}$ defined as

$$x^2 - y^2 = 2(y-x) \Leftrightarrow (x,y) \in S$$

We can find $[0]$ and $[1]$ as follows:

$$[0] = \{x \in \mathbb{R}: xS0\}$$
$$= \{x \in \mathbb{R}: x^2 + 2x = 0^2 + 2(0) = 0\}$$
$$= \{x \in \mathbb{R}: x(x+2) = 0\}$$
$$= \{0,-2\}.$$
$$[1] = \{x \in R: x^2 + 2x = (1)^2 + 2(1) = 3\}$$
$$= \{x \in R: x^2 + 2x - 3 = 0\}$$
$$= \{1,-3\}.\blacksquare$$

Example

We find the equivalence classes for congruence modulo 5 as follows:

Let $a \in Z$

$[a] = \{x \in Z : x \equiv a \text{ (mod5)}\}$

$= \{x \in Z : 5\backslash x - a\}$

$= \{x \in Z : x - a = 5k, k \in Z\}$

$= \{x \in Z : x = a + 5k, k \in Z\}$

$[0] = \{x \in Z : x = 5k, k \in Z\}$

$= \{\dots, -10, -5, 0, 5, 10, \dots.\}$

$[1] = \{x \in Z : x = 1 + 5k, k \in Z\}$

$= \{\dots, -9, -4, 1, 6, 11, \dots\}$

$[2] = \{x \in Z : x = 2 + 5k, k \in Z\}$

$= \{\dots, -8, -3, 2, 7, 12, \dots\}$

$[3] = \cdots.$

$[4] = \cdots$■

In the following section we will discuss two alternative methods for representing relations. One method uses zero-one matrices. The other method uses pictorial representations called directed graphs.

●Representing Relations Using Matrices

A relation between finite sets can be represented using a zero-one matrix. Suppose that R is a relation from $A = \{a_1, a_2, \dots, a_m\}$ to $B = \{b_1, b_2, \dots, b_n\}$. The relation R can be represented by the matrix $M_R = [m_{ij}]$, where

$$m_{ij} = \begin{cases} 1 \text{ if } (a_i, b_j) \in R \\ 0 \text{ if } (a_i, b_j) \notin R \end{cases}$$

Example.

Suppose that $A = \{1, 2, 3\}$ and $B = \{1, 2\}$. Let R be the relation from A to B containing (a, b) if $a \in A$, $b \in B$, and $a > b$. What is the matrix representing R if $a_1 = 1$, $a_2 = 2$, and $a_3 = 3$, and $b_1 = 1$ and $b_2 = 2$?

Solution

Because $R = \{(2,1), (3,1), (3,2)\}$.

The matrix representing R is M_R, where

$$M_R = \begin{bmatrix} m_{11} & m_{12} \\ m_{21} & m_{22} \\ m_{31} & m_{32} \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 1 & 0 \\ 1 & 1 \end{bmatrix}.$$

The 1s in M_R show that the pairs (2, 1), (3, 1), and (3, 2) belong to R. The 0s show that no other pairs belong to R. ■

Example

Let $A = \{a_1, a_2, a_3\}$ and $B = \{b_1, b_2, b_3, b_4, b_5\}$.

The ordered pairs in the relation R represented by the matrix

$M_R = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 0 & 1 \end{bmatrix}$ is:

$R =$

$\{(a_1, b_2), (a_2, b_1), (a_2, b_3), (a_2, b_4), (a_3, b_1), (a_3, b_3), (a_3, b_5)\}$. ■

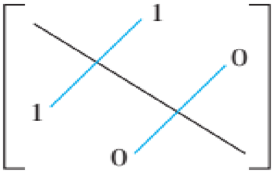
The matrix of a relation on a set, which is a square matrix, can be used to determine whether the relation has certain properties. R is reflexive if and only if $(a_i, a_i) \in R$ for $i = 1, \dots, n$, where $A = \{a_1, \dots, a_n\}$ is the set on which the relation R defined.

Hence R is reflexive if and only if $m_{ii} = 1$ for $i = 1, \dots, n$.

The form of the matrix for an reflexive relation is illustrated in the following figure.

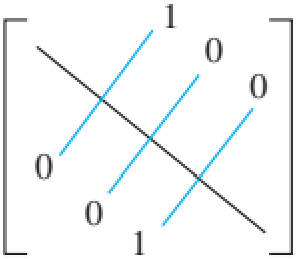
$$\begin{bmatrix} 1 & & & & & & & \\ & 1 & & & & & & \\ & & 1 & & & & & \\ & & & \cdot & & & & \\ & & & & \cdot & & & \\ & & & & & \cdot & & \\ & & & & & & 1 & \\ & & & & & & & 1 \end{bmatrix}$$

The relation R is symmetric if and only if $(a, b) \in R$ implies $(b, a) \in R$. Consequently the relation R on the set $A = \{a_1, a_2, \dots, a_n\}$ is symmetric if and only if $(a_j, a_i) \in R$ whenever $(a_i, a_j) \in R$. Thus R is symmetric if and only if $m_{ji} = 1$ whenever $m_{ij} = 1$. This also means $m_{ji} = 0$ whenever $m_{ij} = 0$. Consequently R is symmetric if and only if $m_{ij} = m_{ji}$ for all pairs i, j with $i = 1, \dots, n$ and $j = 1, \dots, n$. The form of the matrix for an symmetric relation is illustrated in Figure (a).



(a) Symmetric

The relation R is antisymmetric if and only if $(a, b) \in R$ and $(b, a) \in R$ imply that $a = b$. Consequently, the matrix of an antisymmetric relation has the property that if $m_{ij} = 1$ with $i \neq j$, then $m_{ji} = 0$. Or, in other words, either $m_{ij} = 0$ or $m_{ji} = 0$ when $i \neq j$. The form of the matrix for an antisymmetric relation is illustrated in Figure (b).



(b) Antisymmetric

Example

Suppose that the relation R on a set is represented by the matrix

$$M_R = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & 1 \end{bmatrix}.$$

Since all diagonal elements of this matrix are equal to 1, R is reflexive. Moreover, M_R is symmetric $M_R = (M_R)^T$, it follows that R is symmetric. It is easy to see that R is not antisymmetric. ■

Definition

Let $A = [a_{ij}]$ and $B = [b_{ij}]$ be $n \times n$ zero-one matrices. Then the **join** of *A* and *B*, denoted by $A \vee B$, is the zero-one matrix with (*i*, *j*)th entry $a_{ij} \vee b_{ij}$.

The **meet** of *A* and *B*, denoted by $A \wedge B$, is the zero-one matrix with (*i*, *j*)th entry $a_{ij} \wedge b_{ij}$.

Definition.

Let $A = [a_{ij}]$ be an $m \times k$ zero-one matrix and $B = [b_{ij}]$ be $k \times n$ zero-one matrix. Then the **Boolean product** of *A* and *B*, denoted by$A \otimes B$ is the $m \times n$ matrix with (*i*, *j*)th entry $[c_{ij}]$, where $c_{ij} = (a_{i1} \wedge b_{1j}) \vee (a_{i2} \wedge b_{2j}) \vee \ldots \vee (a_{ik} \wedge b_{kj})$.

Example.

The join and meet of the zero-one matrices of *A* and *B*, where $A = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}, B = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 1 & 0 \end{bmatrix}$ are $A \vee B$ and $A \wedge B$ and given as follows:

$$A \vee B = \begin{bmatrix} 1 \vee 0 & 0 \vee 1 & 1 \vee 0 \\ 0 \vee 1 & 1 \vee 1 & 0 \vee 0 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \end{bmatrix}$$

and

$$A \wedge B = \begin{bmatrix} 1 \wedge 0 & 0 \wedge 1 & 1 \wedge 0 \\ 0 \wedge 1 & 1 \wedge 1 & 0 \wedge 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}. \blacksquare$$

Example

The Boolean product $A \otimes B$ of *A* and *B*, where

$$A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 0 \end{bmatrix}, B = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix} \text{ is:}$$

$$A \otimes B = \begin{bmatrix} (1 \wedge 1) \vee (0 \wedge 0) & (1 \wedge 1) \vee (0 \wedge 1) & (1 \wedge 0) \vee (0 \wedge 1) \\ (0 \wedge 1) \vee (1 \wedge 0) & (0 \wedge 1) \vee (1 \wedge 1) & (0 \wedge 0) \vee (1 \wedge 1) \\ (1 \wedge 1) \vee (0 \wedge 0) & (1 \wedge 1) \vee (0 \wedge 1) & (1 \wedge 0) \vee (0 \wedge 1) \end{bmatrix}$$

$$= \begin{bmatrix} 1 \vee 0 & 1 \vee 0 & 0 \vee 0 \\ 0 \vee 0 & 0 \vee 1 & 0 \vee 1 \\ 1 \vee 0 & 1 \vee 0 & 0 \vee 0 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 1 & 0 \end{bmatrix}. \blacksquare$$

The Boolean operations join and meet can be used to find the matrices representing the union and the intersection of two relations as follows:

$M_{R_1 \cup R_2} = M_{R_1} \vee M_{R_2}$and $M_{R_1 \cap R_2} = M_{R_1} \wedge M_{R_2}$

Example.

Suppose that the relations R_1 and R_2on a set A are represented by the matrices

$$M_{R_1} = \begin{bmatrix} 1 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \text{ and } M_{R_2} = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 0 & 0 \end{bmatrix}$$

The matrices representing $R_1 \cup R_2$ and $R_1 \cap R_2$ are

$$M_{R_1 \cup R_2} = M_{R_1} \vee M_{R_2} = \begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 0 \end{bmatrix}$$

$$M_{R_1 \cap R_2} = M_{R_1} \wedge M_{R_2} = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}. \blacksquare$$

We now turn our attention to determining the matrix for the composite of relations. This matrix can be found using the Boolean product of the matrices as follows: $M_{S \circ R} = M_R \otimes M_S$

Example

The matrix representing the relation $S \circ R$ where the matrices representing R and S are

$$M_R = \begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \text{ and } M_S = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 1 \end{bmatrix}:$$

is given as follows:

$$M_{S \circ R} = M_R \otimes M_S = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix}. \blacksquare$$

The matrix representing the **composite** of two relations can be used to find the matrix for M_{R^n}. In particular, $M_{R^n} = M_R^{[n]}$.

Example

Find the matrix representing the relation R^2, where the matrix

representing R is $M_R = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 0 \end{bmatrix}$

Solution

The matrix for R^2 is $M_{R^2} = M_R^{[2]} = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 1 & 1 \\ 0 & 1 & 0 \end{bmatrix}$. ■

● **Representing Relations Using Digraphs**

There is another way of representing a relation on a set using a pictorial representation. Each element of the set is represented by a point and each ordered pair is represented using an arc with its direction indicated by an arrow.

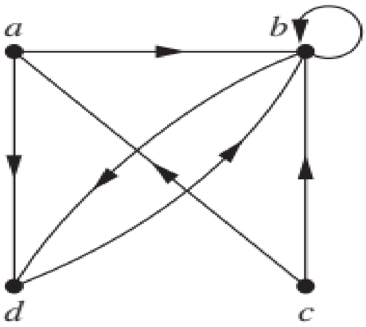
●A directed graph

Definition

A directed graph, or **digraph**, consists of a set of vertices *V* (or nodes) together with a set *E* of ordered pairs of elements of *V* called edges (or arcs). The vertex *a* is called the initial vertex of the edge (*a*, *b*) and the vertex *b* is called the terminal vertex of this edge. An edge of the form (*a*, *a*) is represented using an arc from the vertex a back to itself. It is called a **loop**.◄

Example

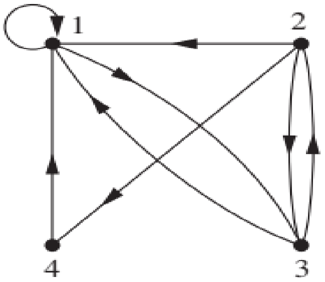
The directed graph with vertices $a, b,$ c, and d, and edges (a, b), (a, d), $(b, b), (b, d), (c, a), (c, b)$, and (d, b) is displayed in the given figure. ■



The relation R on a set A is represented by the directed graph that has the elements of A as its vertices and the ordered pairs (a, b), where $(a, b) \in R$, as edges. This assignment sets up a one-to-one correspondence between the relations on a set A and the directed graphs with A as their set of vertices. Thus, every statement about relations corresponds to a statement about directed graphs, and vice versa. Directed graphs give a visual display of information about relations. As such, they are often used to study relations and their properties. (Note that relations from a set A to a set B can be represented by a directed graph where there is a vertex for each element of A and a vertex for each element of B, as shown above. However, when $A = B$, such representation provides much less insight than the digraph representations described here.) The use of directed graphs to represent relations on a set is illustrated in the following examples.

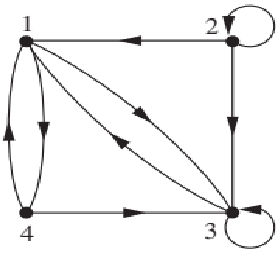
Example

The directed graph of the relation $R = \{(1,1),(1,3),(2,1),(2,3),(2,4),(3,1),(3,2),(4,1)\}$ on the set $\{1,2,3,4\}$ is shown in the given figure. ■



Example

What are the ordered pairs in the relation R represented by the directed graph shown in the given figure?



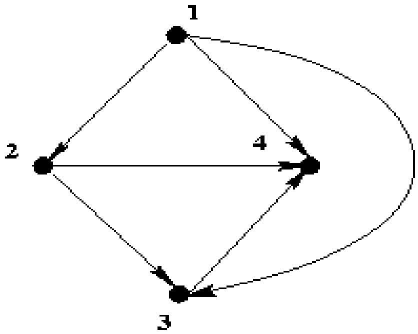
Solution

The ordered pairs (x, y) in the relation are

$R = \{(1,3),(1,4),(2,1),(2,2),(2,3),(3,1),(3,3),(4,1),(4,3)\}$. Each of these pairs corresponds to an edge of the directed graph, with $(2,2)$ and $(3,3)$ corresponding to loops. ■

Example

The less than relation R on the set of integers $A = \{1,2,3,4\}$ is $R = \{(1,2),(1,3),(1,4),(2,3),(2,4),(3,4)\}$ and it can be represented by the given digraph. ■

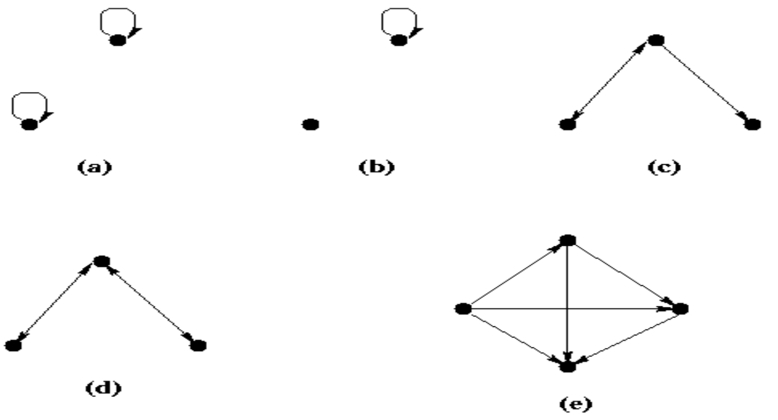


♣The directed graph representing a relation can be used to determine whether the relation has various properties. For instance, a relation is **reflexive** if and only if there is a **loop** at every vertex of the directed graph, so that every ordered pair of the form (x,x) occurs in the relation.

A relation is **symmetric** if and only if for every edge between distinct vertices in its digraph there is an edge in the opposite direction, so that (y,x) is in the relation whenever (x,y) is in the relation. Similarly, a relation is **antisymmetric** if and only if there are never two edges in opposite directions between distinct vertices. Finally, a relation is **transitive** if and only if whenever there is an edge from a vertex x to a vertex y and an edge from a vertex y to a vertex z, there is an edge from x to z (completing a triangle where each side is a directed edge with the correct direction).

Example

The following figures show the digraph of relations with different properties.



(a) is reflexive, antisymmetric, symmetric and transitive.

(b) is not reflexive, and it is antisymmetric, symmetric and transitive.

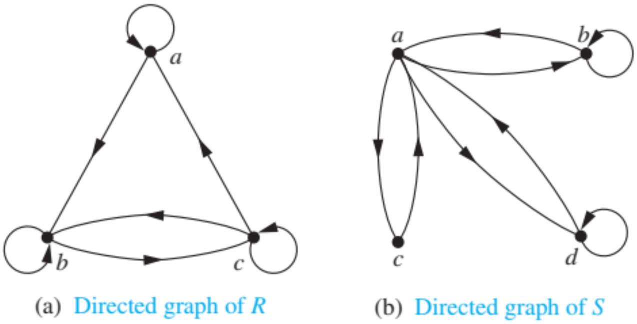
(c) has none of the four properties.

(d) symmetric, but none of the other three.

(e) is antisymmetric and transitive but neither reflexive nor symmetric.■

Example

Determine whether the relations for the directed graphs shown in the following figure are reflexive, symmetric, antisymmetric, and/or transitive.



(a) Directed graph of R (b) Directed graph of S

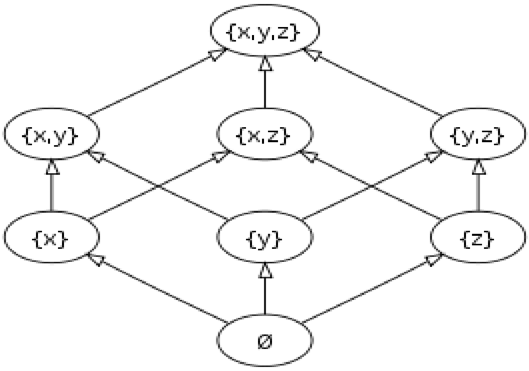
Solution

Because there are loops at every vertex of the directed graph of R, it is reflexive. R is neither symmetric nor antisymmetric because there is an edge from a to b but not one from b to a, but there are edges in both directions connecting b and c. Finally, R is not transitive because there is an edge from a to b and an edge from b to c, but no edge from a to c.

Because loops are not present at all the vertices of the directed graph of S, this relation is not reflexive. It is symmetric and not antisymmetric, because every edge between distinct vertices is accompanied by an edge in the opposite direction. It is also not hard to see from the directed graph that S is not transitive, because (c, a) and (a, b) belong to S, but (c, b) does not belong to S. ■

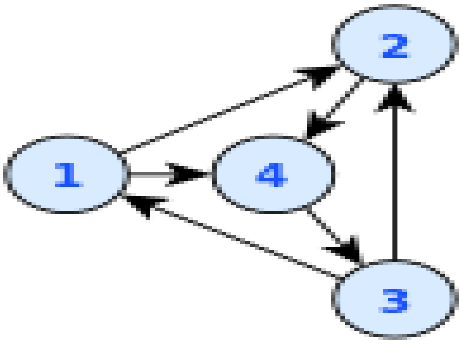
Example

Consider the set $S = \{1, 2, 3\}$. We construct the Hasse diagrams of the partial order $\subseteq$ among the subsets of S as follows:



Example

The ordered pairs in the relation R represented by the directed graph shown in the following figure:



are $R \ = \ \{\,(1,2), (1,4), (2,4), (3,2), (3,1), (4,3)\}$.■

Exercises Set (1.3)

1- List the ordered pairs in the relation R from $A = \{0, 1, 2, 3, 4\}$ to $B = \{0, 1, 2, 3\}$, where $(a, b) \in R$ if and only if

(a) $a = b$; (b) $a + b = 4$; (c) $a > b$; (d) $a \mid b$.

2- For each of these relations on the set $\{1, 2, 3, 4\}$, decide whether it is reflexive, symmetric, antisymmetric or transitive.

(a) $\{(2,2), (2,3), (2,4), (3,2), (3,3), (3,4)\}$;

(b) $\{(1,1), (1,2), (2,1), (2,2), (3,3), (4,4)\}$;

(c) $\{(1,2), (2,3), (3,4)\}$;

(d) $\{(1,1), (2,2), (3,3), (4,4)\}$.

3- Determine whether the relation R on the set of all people is reflexive, symmetric, antisymmetric or transitive, where $(a, b) \in R$ if and only if

(a) a is taller than b;

(b) a and b were born on the same day;

(c) a has the same first name as b;

(d) a and b have a common grandparent.

4- Determine whether the relation R on the set of all integers is reflexive, symmetric, antisymmetric or transitive, where $(a, b) \in R$ if and only if

(a) $x \neq y$; (b) $xy \geq 1$; (c) $x = y + 1$ or $x = y - 1$;

(d) $x \equiv y (\text{mod } 7)$; (e) x is a multiple of y;

(f) x and y are both negative or both nonnegative;

(g) $x = y^2$; (h) $x \geq y^2$.

5- Consider these relations on the set of real numbers

$R_1 = \{(a,b) \in R^2 : a > b\}$;

$R_2 = \{(a,b) \in R^2 : a \geq b\}$;

$R_3 = \{(a,b) \in R^2 : a < b\}$;

$R_4 = \{(a,b) \in R^2 : a \leq b\}$;

$R_5 = \{(a,b) \in R^2 : a \neq b\}$;

$R_6 = \{(a,b) \in R^2 : a \neq b\}$.

Find:

(a) $R_1 \cup R_3,\ R_1 \cup R_5,\ R_2 \cap R_4,\ R_1 - R_2$;

(b) $R_2 \cup R_4,\ R_3 \cup R_6,\ R_3 \cap R_6,\ R_6 - R_3,\ R_2 \oplus R_6$;

(c) $R_1 \circ R_1,\ R_1 \circ R_2,\ R_1 \circ R_3,\ R_2 \circ R_3,\ R_3 \circ R_3$;

(d) $R_2 \circ R_1,\ R_2 \circ R_2,\ R_3 \circ R_6,\ R_5 \circ R_3$.

6- Let R_1 and R_2 be the "divides" and ' is multiple of" relations on the set of all positive integers, respectively.

That is $R_1 = \{(a,b): a \text{ divides } b\}$ and

$R_2 = \{(a,b): a \text{ is a multiple of } b\}$.

Find

(a) $R_1 \cup R_2$; (b) $R_1 \cap R_2$; (c) $R_1 - R_2$;

(d) $R_2 - R_1$; (e) $R_1 \oplus R_2$.

7- Suppose that R and S are reflexive relations on a set A.

Prove or disprove each of these statements

(a) $R \cup S$ is reflexive; b) $R \cap S$ is reflexive;

(c) $R - S$ is reflexive; (d) $R \circ S$ is reflexive.

8- Represent each of these relations on $\{1, 2, 3\}$ with a matrix

(a) $\{(1,1), (1,2), (1,3)\}$;

(b) $\{(1,2), (2,1), (2,2), (3,3)\}$;

(c) $\{(1,1), (1,2), (1,3), (2,2), (2,3), (3,3)\}$;

(d) $\{(1,3), (3,1)\}$.

9- List the ordered pairs in the relation on $\{1,2,3\}$ corresponding to these matrices (where the rows and columns correspond to the integers listed in increasing order).

$$\text{(a) } \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}; \text{ (b) } \begin{bmatrix} 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \end{bmatrix}; \text{ (c) } \begin{bmatrix} 1 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 1 \end{bmatrix}.$$

10- Let R_1, R_2 be relations on a set A represented by the matrices

$$M_{R_1} = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 1 & 1 \\ 1 & 0 & 0 \end{bmatrix} \text{ and } M_{R_2} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}.$$

Find the matrices that represent

(a) $R_1 \cup R_2$; (b) $R_1 \cap R_2$; (c) $R_2 \circ R_1$; (d) $R_1 \circ R_1$; (e)$R_1 \oplus R_2$.

11- Represent each of these relations on $\{1, 2, 3, 4\}$ with a matrix (with the elements of this set listed in increasing order.)

(a) $\{(1,2), (1,3), (1,4), (2,3), (2,4), (3,4)\}$;

(b) $\{(1,1), (1,4), (2,2), (3,3), (4,1)\}$;

(c) $\{(1,2), (1,3), (1,4), (2,1), (2,3), (2,4), (3,1), (3,2)\}$;

(d) $\{(2,4), (3,1), (3,2), (3,4)\}$.

12- List the ordered pairs in the relations on {1,2,3,4} corresponding to these matrices (where the rows and columns correspond to the integers listed in increasing order.)

(a) $\begin{bmatrix} 1 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 \end{bmatrix}$; (b) $\begin{bmatrix} 1 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 1 \end{bmatrix}$;

(c) $\begin{bmatrix} 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \end{bmatrix}$.

13- Draw the digraph representing each of the relations from Exercises 2 and 5.

14- Draw the digraph represents the relation

$\{(a,a),(a,b),(b,c),(c,b),(c,d),(d,a),(d,b)\}$.

15- Picture the divisibility relation on $\{1, 2, \ldots, 12\}$ by a digraph.

16- Determine whether each of the following relations are reflexive, symmetric and transitive:

(i) Relation R in the set $A = \{1, 2, 3 \ldots 13, 14\}$ defined as

$$R = \{(x, y) : 3x - y = 0\};$$

(ii) Relation R in the set $\mathbb{N}$ of natural numbers defined as

$$R = \{(x, y) : y = x + 5 \text{ and } x < 4\};$$

(iii) Relation R in the set $A = \{1, 2, 3, 4, 5, 6\}$ as

$$R = \{(x, y) : y \text{ is divisible by } x\}.;$$

(iv) Relation R in the set $\mathbb{Z}$ of all integers defined as

$$R = \{(x, y) : x - y \text{ is as integer}\};$$

(v) Relation R in the set *A* of human beings in a town at a particular time given by

(a) R = {(*x*, *y*): *x* and *y* work at the same place};

(b) R = {(*x*, *y*): *x* and *y* live in the same locality};

(c) R = {(*x*, *y*): *x* is exactly 7 cm taller than *y*};

(d) R = {(*x*, *y*): *x* is wife of *y*};

(e) R = {(*x*, *y*): *x* is father of *y*};

17- Show that the relation R in the set $\mathbb{R}$ of real numbers, defined as $R = \{(a, b): a \leq b^2\}$ is neither reflexive nor symmetric nor transitive.

18- Check whether the relation R defined in the set $\{1, 2, 3, 4, 5, 6\}$ as $R = \{(a, b): b = a + 1\}$ is reflexive, symmetric or transitive.

19- Show that the relation R in $\mathbb{R}$ defined as $R = \{(a, b): a \leq b\}$, is reflexive and transitive but not symmetric.

20- Check whether the relation R in $\mathbb{R}$ defined as $R = \{(a, b): a \leq b^3\}$ is reflexive, symmetric or transitive.

21- Show that the relation R in the set $\{1, 2, 3\}$ given by $R = \{(1, 2), (2, 1)\}$ is symmetric **but** neither reflexive nor transitive.

22- Show that the relation R in the set $A = \{1, 2, 3, 4, 5\}$ given by $R = \{(a, b): |a - b| \text{ is even}\}$, is an equivalence relation. Show that all the elements of $\{1, 3, 5\}$ are related to each other and all the elements of $\{2, 4\}$ are related to each other. But no element of $\{1, 3, 5\}$ is related to any element of $\{2, 4\}$.

23. Given an example of a relation. Which is

(i) Symmetric but neither reflexive nor transitive.

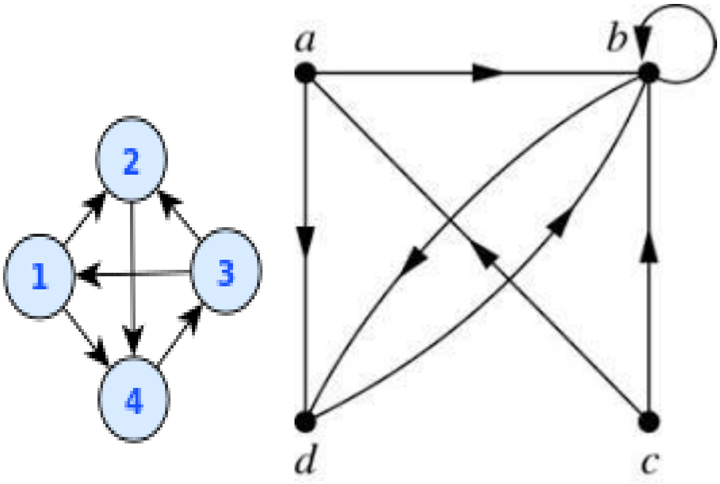
(ii) Transitive but neither reflexive nor symmetric.

(iii) Reflexive and symmetric but not transitive.

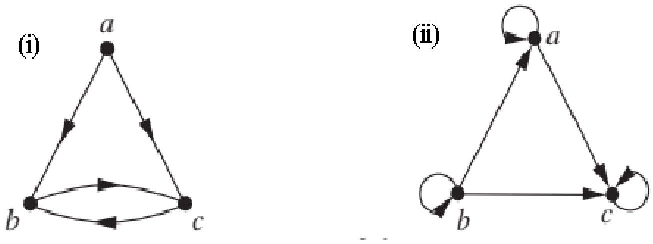
(iv) Reflexive and transitive but not symmetric.

(v) Symmetric and transitive but not reflexive.

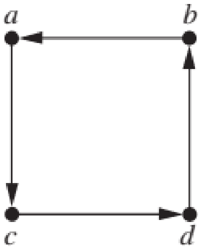
24- What are the ordered pairs in the relation *R* represented by the directed graph shown in the following figure?



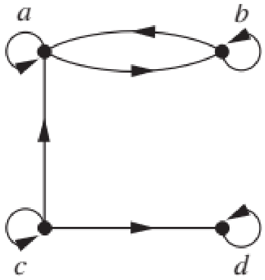
25- What are the ordered pairs in the relation *R* represented by the directed graph shown in the following figure?



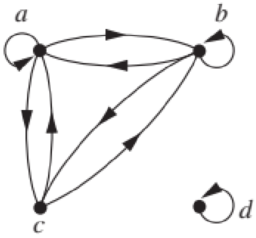
(iii)



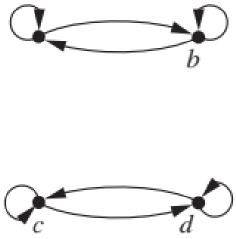
(iv)



(v)



(vi)



26. Given the directed graphs representing two relations, how can the directed graph of the union, intersection, symmetric difference, difference, and composition of these relations be found?

CHAPTER (II)

MATHEMATICAL LOGIC

Chapter (III)

Mathematical Logic

The rules of mathematical logic specify methods of reasoning mathematical statements. Greek philosopher, Aristotle, was the pioneer of logical reasoning. Logical reasoning provides the theoretical base for many areas of mathematics and consequently computer science. It has many practical applications in computer science like design of computing machines, artificial intelligence, definition of data structures for programming languages etc.

2.1 Propositional Calculus

Propositional Logic is concerned with statements to which the truth values, "true" and "false", can be assigned. The purpose is to analyze these statements either individually or in a composite manner.

Definition.

In logic, **a proposition (or a statement)** is a meaningful declarative sentence that is either true or false, but not both.

The truth value of a proposition is True "T or 1" if it is a true proposition and false "F or 0" if it is a false proposition. Letters p, q, r , ... are used to denote proposition and are called propositional variables.

* The following propositions are true

(i) A triangle has three sides.

(ii) 7 is odd.

(iii) 2 divides 24.

* The following propositions are false:

(i) $5 + 3 = 9$.

(ii) Makkah is the capital of Saudi Arabia.

(iii) 2 divides 7.

* The following are not proposition:

(1) Who are you?

Not declarative sentences

(2) Help yourself!

Not declarative sentence.

(3) $u - 2 = 1$

Neither true nor false.

(4) $u - v = w$.

Neither true nor false.

(5) Broccoli tastes good.

Meaningful declarative sentences, but is not proposition but rather matters of opinion or taste.

Definition.

A formula (or a compound proposition) A formula is formed from existing propositions using connectives.

Definition.

Since we need to know the truth value of a proposition in all possible scenarios, we consider all the possible combinations of the propositions which are joined together by Logical Connectives to form the given compound proposition. This compilation of all possible scenarios in a tabular format is called a **truth table**.

In particular, truth tables can be used to tell whether a propositional expression is true or false for all legitimate input values. Practically, a truth table is composed of one column for each input variable (for example, p and q), and one final column for all of the possible results of the logical operation that the table is meant to represent (for example, $p \rightarrow q$). Each row of the truth table therefore contains one possible configuration of the input variables (for instance, p is true (written 1 or T) q is false (written 0 or F)), and the result of the operation for those values.

♣ Logical Connectives

Connectives are either unary operations like logical identity and logical negation, or binary operations like logical conjunction, logical disjunction and logical implication.

Definition. (Logical identity and logical Negation).

Let p be a proposition.

● Logical identity

Logical identity is an operation on one logical value, typically the value of a proposition that produces a value of *true* if its operand is true and a value of *false* if its operand is false. The truth table for the logical identity operator is as follows:

Logical Identity

p	p
Operand	*Value*
1	1
0	0

● Logical negation

Logical negation is an operation on one logical value, typically the value of a proposition, which produces a value of *true* if its operand is false and a value of *false* if its operand is true.

The truth table for logical negation (written as ¬*p* or ~*p*) is as follows:

Logical negation

p	$\neg p$
1	0
0	1

Example.

The negation of the proposition "The sun shines on the screen" is "The sun does not shine on the screen".■

We will now introduce the logical connectives (binary operations) that are used to form formulas.

Definition. (Logical Conjunction" ∧ ")

Logical conjunction is an operation on two logical values, typically the values of two propositions, that produces a value of *true* if both of its operands are true.

The truth table for p AND q (written as $p \wedge q$) is as follows:

Logical Conjunction

p	q	$p \wedge q$
1	1	1
1	0	0
0	1	0
0	0	0

Example.

Let p be the proposition “It is sunny today” and q be the proposition “The sun shines on the screen”. Then the conjunction of these propositions, $p \wedge q$, is the proposition “It is sunny today and the sun shines on the screen”. This proposition is true when the day is sunny and the sun shines on the screen. It is false otherwise. ■

Definition. (Logical Disjunction" ∨ ")

Logical disjunction is an operation on two logical values, typically the values of two propositions, that produces a value of *true* if at least one of its operands is true.

The truth table for p OR q (written as $p \vee q$) is as follows:

Logical disjunction

p	q	$p \vee q$
1	1	1
1	0	1
0	1	1
0	0	0

Example.

The disjunction of the propositions p and q where p and q are the same propositions as in the above example, $p \vee q$, is the proposition “It is sunny today or the sun shines on the screen”. This proposition is true on any day that is either sunny day or the sun shines on the screen (including both). It is only false on days that are not sunny and when it also does not shine on the screen. ■

Definition.

(“Logical Implication” or “Conditional Statement” " → ")

Logical implication is associated with an operation on two logical values, typically the values of two propositions, that produces a value of *false* just in the singular case the first operand is true and the second operand is false.

The truth table associated with the Logical implication if p then q (symbolized as $p \rightarrow q$) is as

Logical implication

p	q	$p \longrightarrow q$
1	1	1
1	0	0
0	1	1
0	0	1

It may also be useful to note that $p \rightarrow q$ and $\neg p \vee q$ have the same truth table. A variety of terminology is used to express $p \rightarrow q$. Some of them are: “if *p*, then *q*”, “*p* implies *q*”, “if *p*, *q*” , “*p* only if *q*”, “*p* is sufficient for *q*”, “a sufficient condition for *q* is *p*”, “*q* if *p*”, “*q* whenever *p*”, “*q* when *p*” , “*q* is necessary for *p*”
“a necessary condition for *p* is *q*” , “*q* follows from *p*”
and "*q* unless $\neg p$.

Example.

Let p the proposition "Aly study well" and q the proposition "Aly will be a Computer Science student". Then the formula $p \rightarrow q$ -as a formula in English- is "If Aly study well, then he will be a Computer Science student ". ■

Definition. (Converse, Contra-positive and Inverse)

There are some related conditional statements that can be formed from $p \to q$. The conditional statement $q \to p$ is called the **converse** of $p \to q$. The **contra-positive** of $p \to q$ is the conditional statement $\neg q \to \neg p$.

The statement $\neg p \to \neg q$ is called the **inverse** of $p \to q$.

The contra-positive, $\neg q \to \neg p$, of a conditional statement $p \to q$ has the same truth value as $p \to q$.

On the other hand, neither the converse, $q \to p$, nor the inverse $\neg p \to \neg q$, has the same truth value as $p \to q$ for all possible truth values of p and q.

Example.

What are the contra-positive, the converse, and the inverse of the conditional statement "The home team wins whenever it is raining".

Solution.

Because "q whenever p" is one of the ways to express the conditional statement $p \to q$, the original statement can be rewritten as "If it is raining, then the home team wins".

Consequently, the contra-positive of this conditional statement is “If the home team does not win, then it is not raining”.

The converse is “If the home team wins, then it is raining”.

The inverse “If it is not raining, then the home team does not win”. Only the contrapositive is equivalent to the original statement. ■

We now introduce another way to combine propositions.

Definition. (Biconditional " ↔ ").

Biconditional (also known as **logical equality**) is an operation on two logical values, typically the values of two propositions, that produces a value of *true* if both operands are false or both operands are true.

The truth table for p XNOR q (written as $p \leftrightarrow q$) is as follows:

Logical Equality

p	q	$p \leftrightarrow q$
1	1	1
1	0	0
0	1	0
0	0	1

So $p \leftrightarrow q$ is true if p and q have the same truth value (both true or both false), and false if they have different truth values. There are some other ways to express $p \leftrightarrow q$ "p is necessary and sufficient for q"; "p iff q" where "iff" is the abbreviation for "if and only if" and " if p then q and conversely ".

Example.

Let p be the statement “You can pass the exam.” and let q be the statement “You study well”. Then $p \leftrightarrow q$ is the statement “You can pass the exam if and only if you study well”. ■

Remark.

The previous operators (¬, ∧, ∨, →, ↔) are the **common operators** which we will focus on.

Definition. (Exclusive Or" $\oplus$ "**).**

Truth table for Exclusive Or " $\oplus$ "

Logical Equality

p	q	$p \oplus q$
1	1	0
1	0	1
0	1	1
0	0	0

Actually, this operator can be expressed by using other operators:

$p \oplus q$ is the same as $\neg\,(p \ \leftrightarrow\ q)$.

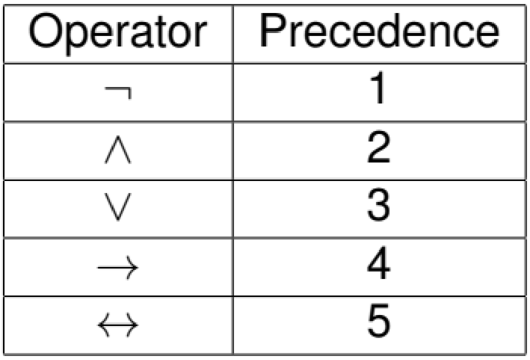
$\oplus$ is used often in CSE. So we have a symbol for it.

● Order of precedence

As a way of reducing the number of necessary parentheses, one may introduce precedence rules for operators. $\neg$ has higher precedence than $\wedge$, $\wedge$ higher than $\vee$, and $\vee$ higher than $\longrightarrow$.

Here is a table that shows a commonly used precedence of logical operators.

The order of precedence determines which connective is the "main connective" when interpreting a formula.



Operator	Precedence
$\neg$	1
$\wedge$	2
$\vee$	3
$\rightarrow$	4
$\leftrightarrow$	5

Example.

$\neg p \wedge q$ means $(\neg p) \wedge q$;

$p \wedge q \longrightarrow r$ means $(p \wedge q) \longrightarrow r$;

$p \vee q \wedge \neg r \rightarrow s$ is short for $\left[p \vee \left(q \wedge (\neg r)\right)\right] \rightarrow s$.

When in doubt, use parenthesis. ■

Example.

Find the truth table for the following formula: "If you studied discrete Mathematics well and did not neglect studying logic, you would gain high marks in the exam".

Solution.

Suppose that

p: studied discrete Mathematics well;

q: neglect studying logic;

r: gain high mark in the exam.

The formula is $p \land \neg q \longrightarrow r$

p	q	r	$\neg q$	$p \land \neg q$	$p \land \neg q \longrightarrow r$
1	1	1	0	0	1
1	1	0	0	0	1
1	0	1	1	1	1
1	0	0	1	1	0
0	1	1	0	0	1
0	1	0	0	0	1
0	0	1	1	0	1
0	0	0	1	0	1

■

● Tautologies and Contradictions

Definition.

A formula that is always true, no matter what the truth values of the propositions that occur in it, is called a **tautology**.

A formula that is always false is called **contradiction**.

A formula that is neither a tautology nor a contradiction is called a **contingency**.

Example.

We can construct examples of tautologies and contradictions using just one proposition. Consider the truth tables of $p \lor \neg p$ and $p \land \neg p$. Since $p \lor \neg p$ is always true, it is a tautology. Since $p \land \neg p$ is always false, it is a contradiction.

Example of a tautology and a contradiction

p	$\neg p$	$p \lor \neg p$	$p \land \neg p$
1	0	1	0
0	1	1	0

■

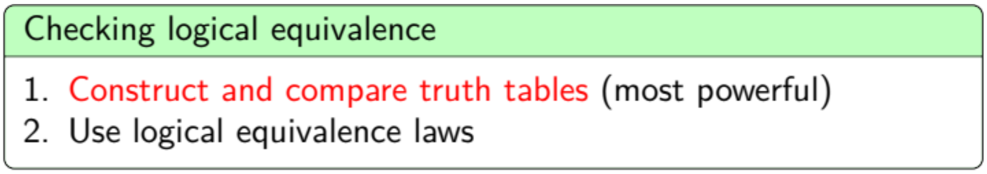
● Logical Equivalence

Definition.

Two formulas p and q are **logically equivalent**, denoted by $p \equiv q$, if and only if they have the same truth values for all possible combination of truth values for the propositional variables. Also,

Definition.

Two formulas p and q are called **logically equivalent** if $p \leftrightarrow q$ is a tautology.



Checking logical equivalence

1. Construct and compare truth tables (most powerful)
2. Use logical equivalence laws

Example.

The formulas $p \rightarrow q$ and $\neg p \vee q$ are logically equivalent.

p	q	$\neg$p	$p \rightarrow q$	$\neg p \vee q$
1	1	0	1	1
1	0	0	0	0
0	1	1	1	1
0	0	1	1	1

Example.

The formulas $\neg(p \vee q)$ and $\neg p \wedge \neg q$ are logically equivalent.

p	q	$\neg \mathrm{p}$	$\neg q$	$p \vee q$	$\neg(p \vee q)$	$\neg p \wedge \neg q$
1	1	0	0	1	0	0
1	0	0	1	1	0	0
0	1	1	0	1	0	0
0	0	1	1	0	1	1

Since the truth values of the formulas $\neg(p \vee q)$ and $\neg p \wedge \neg q$ agree for all possible combinations of the truth values of p and q, it follows that $\neg(p \vee q) \leftrightarrow \neg p \wedge \neg q$ is a tautology and these formulas are logically equivalent. Similarly, we can prove that $\neg(p \wedge q) \equiv \neg p \vee \neg q$.■

Theorem. (Algebraic properties of connectives)

(1) Commutative rules:

(a) $p \wedge q \equiv q \wedge p,$ (b) $p \vee q \equiv q \vee$ p.

(2) Associative rules:

(a) $(p \wedge q) \wedge r \equiv p \wedge (q \wedge r),$

(b) $(p \vee q) \vee r \equiv p \vee (q \vee r).$

(3) Distributive rules:

(a) $p \wedge (q \vee r) \equiv (p \wedge q) \vee (p \wedge r),$

(b) $p \vee (q \wedge r) \equiv (p \vee q) \wedge (p \vee r).$

(4) Identity rules:

(a) $p \vee 0 \equiv p,$ (b) $p \wedge 1 \equiv p$

(5) Negation rules:

$p \wedge \neg p \equiv 0$ and $p \vee \neg p \equiv 1$.

(6) Double negation rule:

$\neg(\neg p) \equiv p.$

(7) Idempotent rules:

$p \vee p \equiv p$ and $p \wedge p \equiv p$.

(8) De Morgan's rules:

(a) $\neg(p \wedge q) \equiv \neg p \vee \neg q$,

(b) $\neg(p \vee q) \equiv \neg p \wedge \neg q.$

(9) Universal rules:

$p \wedge 0 \equiv 0$ and $p \vee 1 = 1.$

(10) Absorption rules:

(a) $p \vee (p \wedge q) \equiv p,$

(b) $p \wedge (p \vee q) \equiv p.$

(11) Alternative proof rule:

(a) $p \longrightarrow (q \vee r) \equiv (p \wedge \neg q) \longrightarrow r \equiv (p \wedge \neg r) \longrightarrow q.$

(b) $p \vee q \longrightarrow r \equiv (p \longrightarrow r) \wedge (q \longrightarrow r).$

(12) Conditional rules:

(a) $p \longrightarrow q \equiv \neg p \vee q$

(b) $\neg(p \longrightarrow q) \equiv p \wedge \neg q.$

(13) Biconditional rules:

(a) $p \leftrightarrow q \equiv (p \longrightarrow q) \wedge (q \longrightarrow p)$

(b) $p \leftrightarrow q \equiv (p \wedge q) \vee (\neg p \wedge \neg q)$

(c) $p \leftrightarrow q \equiv (\neg p \vee q) \wedge (p \vee \neg \mathrm{q})$

(14) Rules of contrapositive:

$$p \longrightarrow q \equiv \neg\, q \longrightarrow \neg p$$

(15) Exportation – importation rule:

$$p \longrightarrow (q \longrightarrow r) \equiv p \wedge q \longrightarrow r$$

Proof. Exercise. ◄

Example.

Use the algebraic properties of connectives to prove:

(a) $\neg\big(p \wedge (\neg p \vee q)\big) \equiv \neg p \vee \neg q$;

(b) $[(p \vee q) \wedge (p \longrightarrow r) \wedge (q \longrightarrow r)] \longrightarrow r$ is a tautology.

Solution.

(a) Exercise.

(b) $[(p \vee \mathrm{q}) \wedge \big((p \longrightarrow r) \wedge (q \longrightarrow r)\big)] \longrightarrow r$

$\equiv [(p \vee q) \wedge \big((p \vee q) \longrightarrow r\big)] \rightarrow r$

Alternative proof rule

$\equiv [(p \vee \mathrm{q}) \wedge (\neg(p \vee q) \vee \mathrm{r})] \longrightarrow r$

Conditional rule

$\equiv [((p \vee q) \wedge (\neg(p \vee q))) \vee \big((p \vee q) \wedge r\big)] \longrightarrow r$

Distributive rule

$\equiv [0 \vee ((p \vee q) \wedge r)] \longrightarrow r$ Negation rule

$\equiv [(p \vee q) \wedge r] \longrightarrow r$ Identity rule

$\equiv \neg[(p \vee q) \wedge r] \vee r$ Conditional rule

$\equiv [\neg(p \vee q) \vee \neg r] \vee r$ De Morgan's rule

$\equiv \neg(p \vee q) \vee [\neg \mathrm{r} \vee r]$ Associative rule

$\equiv \neg(p \vee q) \vee 1$ Negation rule

$\equiv 1$ Idempotent rules. ■

Exercise Set (2.1)

1- Which of the following are propositions?

(a) Buy Premium Bonds!

(b) The Apple Macintosh is a 16-bit computer.

(c) There is a largest even number.

(d) Why are we here?

(e) $8 + 7 = 13$.

(f) $a + b = 13$.

2- p is "1024 bytes is known as 1MB" and q is "A computer keyboard is an example of a data input device". Express the following formulas as English sentences in as natural a way as you can. Are the resulting propositions true or false?

(a) $p \wedge q$; (b)) $p \vee q$; ; (c) $\neg p$.

3- p is "$x < 50$"; q is "$x > 40$".

Write as simply as you can:

(a) $\neg p$; (b) $\neg q$; (c) $p \wedge q$; (d) $p \vee q$; (e) $\neg p \wedge q$;

(f) $\neg p \wedge \neg q$.

One of these compound propositional functions always produces the output *true*, and one always outputs *false*. Which ones?

4- p is "I like Math" and q is "I am going to spend at least 6 hours a week on Math". Write in as simple English as you can:

(a) $(\neg p) \wedge q;$ (b) $(\neg p) \vee q;$

(c) $\neg(\neg p);$ (d) $(\neg p) \vee (\neg q);$

(e) $\neg(p \vee q);$ (f) $(\neg p) \wedge (\neg q);$

(g) $p \to q$; (h) $p \wedge q.$

5- Construct a truth table for each of these formulas:

(a) $p \wedge \neg p;$

(b) $p \vee \neg p;$

(c) $(p \vee \neg q) \to q;$

(d) $(p \vee q) \to (p \wedge q);$

(e) $p \to \neg p;$

(f) $p \leftrightarrow \neg p.$

6- Show that each of these implications is a tautology by using truth tables.

(a) $[\sim p \wedge (p \vee q)] \to q.$

(b) $[(p \to q) \wedge (q \to r)] \wedge (p \to r)$

7- Show that each implication in Exercise 6 is a tautology without using truth tables.

8- Show that every pair in the following are logically equivalent:

(a) $p \to q$ and $\neg q \to \neg p$

(b)$\neg p \leftrightarrow q$and$p \leftrightarrow \neg q$

(c) $\neg(p \leftrightarrow q)$ and$\neg p \leftrightarrow \neg q$

(d) $(p \to q) \wedge (p \to r)$ and$p \to (q \wedge r)$

(e) $(p \to q) \vee (p \to r)$ and$p \to (q \vee r)$

9- Show that$(p \vee q) \wedge (\neg p \vee r) \to (q \vee r)$ is a tautology.

10- Show that $(p \to q) \to r$ and $p \to (q \to r)$ are not logically equivalent.

11-Prove that:

(a) $p \longrightarrow q \equiv \neg q \longrightarrow \neg p$;

(b) $\neg(p \vee q) \equiv \neg p \vee \neg q$;

(c) $p \longrightarrow q \equiv \neg p \vee q$;

(d) $(p \wedge q) \longrightarrow r \equiv \neg r \longrightarrow (\neg p \vee \neg q)$.

2.2 Predicates and Quantifiers

(A) Predicates

Predicates are statements involving variables (called predicate variables), such as:

"$x > 3$", "$x = y + 3$", "$x + y = z$".

They are not propositions because the truth value you give them will depend on the values assigned to the variables x and y. **The domain** of a predicate variable is the set of all values that may be substituted in place of the variable.

In English you may have statements like this:

1- She is Tall and Fair.

2- x was born in a city y in the year z.

Often pronouns (I, he, she, you etc.) are used in place of variables.

In the first case - we cannot say if the statement is true because that depends of who she is and in the second case the statement will get a truth value depending on variable x, y and z.

Predicate are noted something like this $P(x, y, z)$.

For example

$P(x, y, z)$. This stands for the predicate "$x + y = z$".

$M(x, y)$. This stands for "x is married to y".

In general, you have predicates in the form of:

$P(x)$ - this is a unary predicate (has one variable).

$P(x, y)$ - this is a binary predicate (has two variables).

$P(x_1, x_2, \ldots, x_n)$ - this is an n-ray or n-place predicate – (has n individual variables in a predicate).

You have to choose the values for the variables - these can be from a set of humans - a specific human, a set of places or a place, a set of integers or an integer, a set of real numbers or a real number and so on.

The values are chosen from a particular domain of values called **a universe or a universe of discourse**.

If we take a look at this again:

x was born in a city *y* in the year *z*. *x* is taken from a set of human beings, *y* is taken from a set of cities and *z* is taken from a set of years. This is called the underlying universe.

Looking at this again:

$P(x, y, z)$.The values for the variables *x*, *y* and *z* will be taken from a set of integers or negative integers.

In some cases, you will have to specify the underlying universe because a certain predicate may be true for real numbers but false for not real numbers.

In the case *x* has to be a human being and *y* has to be a city and *z* has to be a year. You cannot have *y* as an integer or *z* a colour for example.

If you assign a particular value to each of the n place values in $P(x_1, x_2, \dots, x_n)$ then the **predicate** becomes a **proposition** and takes a truth value - true or false.

***Again** the statement "x is greater than* 3*" has two parts. The first part, the variable x, is the subject of the statement. The second part, the predicate, "is greater than* 3*", refers to a property that the subject of the statement can have. We can denote the statement "x is greater than* 3*" by* $P(x)$ *where P denotes the predicate "is greater than* 3*" and x is the variable. Once a value has been assigned to the variable x, the statement* $P(x)$ *becomes a proposition and has a truth value.*

Example.

Let $P(x)$ denote the statement "$x > 3$". What are the truth values of the propositions $P(4)$ and $P(2)$?

Solution.

We obtain the proposition $P(4)$ by setting $x = 4$ in the statement "$x > 3$". Hence $P(4)$, which is the proposition "$4 > 3$" is true.

However, $P(2)$ which is the proposition"$2 > 3$", is false.■

Example.

Let $Q(x, y)$ denote the statement "$x = y + 3$." What are the truth values of the propositions $Q(1, 2)$ and $Q(3, 0)$?

Solution.

To obtain proposition $Q(1,2)$, set $x = 1$ and $y = 2$ in the statement $Q(x, y)$. Hence$Q(1, 2)$ is the proposition "$1 = 2 + 3$" which is false.

The proposition $Q(3, 0)$ is the proposition "$3 = 0 + 3$" which is true. ■

Example.

What are the truth values of the propositions $P(1,2,3)$ and $P(0,0,1)$, where $P(x,y,z)$ denote the statement "$x + y = z$"?

Solution.

The proposition $P(1,2,3)$ is obtained by setting $x = 1, y = 2$, and $z = 3$ in the statement $P(x,y,z)$. We see that $P(1,2,3)$ is the proposition"$1 + 2 = 3$", which is true.

Also, note that $P(0,0,1)$, which is the proposition"$0 + 0 = 1$" is false.■

Definition.

If $P(x)$ is a predicate and *x* has domain *D*, the **truth set** of $P(x)$ is the set of all elements of *D* that make $P(x)$ true when they are substituted for *x*. The truth set of $P(x)$ is denoted $\{x \in D : P(x)\}$ and we read as "the set of all *x* in *D* such that $P(x)$."

Example.

Let $Q(n)$ be the predicate "n is a factor of 8." Find the truth set of $Q(n)$ if:

(a) the domain of n is $\mathbb{Z}^+$, the set of all positive integers.

(b) the domain of n is $\mathbb{Z}$, the set of all integers.

Solution.

(a) The truth set is $\{1, 2, 4, 8\}$ because these are exactly the positive integers that divide 8 evenly.

(b) The truth set is $\{1,\ 2,\ 4,\ 8, -1, -2, -4, -8\}$ because the negative integers $-1, -2, -4$, and -8 also divide into 8 without leaving a remainder. ■

Definition.

Let $P(x)$ and $Q(x)$ be predicates with common domain D of x. The notation $P(x) \Rightarrow Q(x)$ means that every element in the truth set of $P(x)$ is in the truth set of $Q(x)$. Similarly, $P(x) \Leftrightarrow Q(x)$ means that $P(x)$ and $Q(x)$ have identical truth sets.

Example.

Let $P(x)$ be "x is a factor of 8",

$Q(x)$ be "x is a factor of 4",

$R(x)$ be " $x < 5$ and $x \neq 3$",

and let the domain of x be set of positive integers. Then

Truth set of $P(x)$ is $\{1, 2, 4, 8\}$.

Truth set of $Q(x)$ is $\{1, 2, 4\}$.

Since every element in the truth set of $Q(x)$ is in the truth set of $P(x)$, then $Q(x) \Rightarrow P(x)$.

Further, truth set of $R(x)$ is $\{1, 2, 4\}$, which is identical to the truth set of $Q(x)$. Hence $R(x) \Leftrightarrow Q(x)$. ■

(B) Quantifiers

(i) The Universal Quantifier " ∀ "

One sure way to change predicates into propositions is to assign specific values to all their variables.

For example, if *x* represents the number 35, the sentence "*x* is divisible by 5" is a true proposition.

Another way to obtain propositions from predicates is to add quantifiers. Quantifiers are words that refer to quantities such as "some" or "all" and tell for how many elements a given predicate is true.

The symbol ∀ is called the universal quantifier.

Depending on the context, it is read as "for every," "for each," "for any," "given any," or "for all."

For example, another way to express the sentence

"Every human being is mortal"

or

"All human beings are mortal"

is to write

"∀ human beings *x*, *x* is mortal",

which you would read as

“For every human being *x*, *x* is mortal.”

If you let *D* be the set of all human beings, then you can symbolize the statement more formally by writing

“$\forall x \in D$, *x* is mortal”.

In sentences containing a mixture of symbols and words, the ∀ symbol can refer to two or more variables.

For instance, you could symbolize

“For all real numbers *x* and *y*, $x + y = \; y + x$.”

as

“∀ real numbers *x* and *y*, $x + y = \; y + x$.”

Definition.

Let $P(x)$ be a predicate and D the domain of x. A universal quantification of $P(x)$ is a proposition of the form “$\forall x \in D, P(x)$.” It is defined to be true if, and only if, $P(x)$ is true for each individual *x* in D. It is defined to be false if, and only if, $P(x)$ is false for at least one *x* in *D*.

The notation $\forall x P(x)$ is used for the universal quantification of $P(x)$ when the domain is known.

Here ∀ is called the **universal quantifier**.

Example.

Let $P(x)$ be the statement "$x + 1 > x$". What is the truth value of the quantification $\forall x P(x)$, where the domain consists of all real numbers?

Solution.

Since $P(x)$ is true for all real numbers x, the quantification $\forall x P(x)$ is true. ■

Example.

Let $Q(x)$ be the statement "$x < 2$". What is the truth value of the quantification $\forall x Q(x)$, where the domain consists of all real numbers?

Solution.

$Q(x)$ is not true for every real number x, since, for instance, $Q(3)$ is false. Thus $\forall x Q(x)$is false. ■

Note.

When all the elements in the universe of discourse can be listed, say $x_1, x_2, \ldots, x_n$ it follows that the universal quantification $\forall x P(x)$ is the same as the conjunction $P(x_1) \wedge P(x_2) \wedge \ldots \wedge P(x_n)$.

Example.

What is the truth value of $\forall x P(x)$, where $P(x)$ is the statement "$x^2 < 10$" and the universe of discourse consists of positive integers not exceeding 4?

Solution.

The statement $\forall x P(x)$ is the same as the conjunction $P(1) \wedge P(2) \wedge P(3) \wedge P(4)$. Since $P(4)$,which is the statement"$4^2 < 10$" , is false, so $\forall x P(x)$ is false. ■

To show that a statement of the form $\forall x P(x)$ is false, where $P(x)$ is a propositional function, we need only find one value of x in the universe of discourse for which $P(x)$ is false. Such a value of x is called a **counterexample** to the statement $\forall x P(x)$.

Example.

Suppose that $P(x)$ is "$x^2 > 0$". To show the statement $\forall x P(x)$ is false where the universe of discourse consists of all integers, we give a counterexample. We see that $x = 0$ is a counterexample since $x^2 = 0$ when $x = 0$ so that x^2 is not greater than 0 when $x = 0$. ■

(ii) The Existential Quantifier " ∃ "

The symbol ∃ denotes "there exists" and is said to be the existential quantifier. For example, the sentence

"There is a student in Math211"

can be written as

"∃ a person *x* such that *x* is a student in Math211",

or, more formally,

"$\exists x \in P$ such that *x* is a student in Math211",

where P is the set of all people.

The domain of the predicate variable is generally indicated either between the ∃ symbol and the variable name or immediately following the variable name, and the words such that are inserted just before the predicate. Some other expressions that can be used in place of there exists are there is a, we can find a, there is at least one, for some, and for at least one.

In a sentence such as

"∃ integers *m* and *n* such that $m + n = m \cdot n$,"

the ∃ symbol is understood to refer to both *m* and *n*.

In more formal versions of symbolic logic, the words such that are not written out (although they are understood) and a separate ∃ symbol is used for each variable: "$\exists m \in \mathbb{Z}\, (\exists n \in \mathbb{Z}(m+n = m \cdot n))$."

Definition.

Let $P(x)$ be a predicate and D the domain of x. An existential statement is a statement of the form

"$\exists x \in D$ such that $P(x)$."

It is defined to be true if, and only if, $P(x)$ is true for at least one x in D. It is false if, and only if, $P(x)$ is false for all x in D.

We use the notation $\exists x P(x)$ for the existential quantification of $P(x)$.

Here ∃ is called the **existential quantifier**.

A domain must always be specified when a statement $\exists x P(x)$ is used. Furthermore, the meaning of $\exists x P(x)$ changes when the domain changes. Without specifying the domain, the statement $\exists x P(x)$ has no meaning. The existential quantification $\exists x P(x)$ is read as:

"There is an x such that $P(x)$","There is at least one x such that $P(x)$" or "For some x $P(x)$".

Example.

Let $P(x)$ denote the statement “$x > 3$”. What is the truth value of the quantification $\exists xP(x)$, where the domain consists of all real numbers?

Solution.

Because “$x > 3$” is sometimes true - for instance, when $x = 4$, the existential quantification $\exists xP(x)$ of $P(x)$ is true.■

Example.

Let $Q(x)$ denote the statement "$x = x + 1$". What is the truth value of the quantification $\exists xP(x)$, where the domain consists of all real numbers?

Solution.

Because $Q(x)$ is false for every real number x, the existential quantification of $Q(x)$ which is $\exists xP(x)$ is false.■

When all elements in the domain can be listed say $x_1, x_2, \ldots, x_n$ the existential quantification $\exists xP(x)$ is the same as the disjunction $P(x_1) \vee P(x_2) \vee \ldots \vee P(x_n)$

because this disjunction is true if and only if at least $P(x_1), P(x_2), \dots, P(x_n)$ is true.

Example.

What is the truth value of $\exists x P(x)$, where $P(x)$ is the statement "$x^2 > 10$" and the domain consists of the positive integers not exceeding 4?

Solution.

As the domain is $\{1, 2, 3, 4\}$, the proposition $\exists x P(x)$ is the disjunction $P(1) \lor P(2) \lor P(3) \lor P(4)$.

Because $P(4)$, which is the statement "$4^2 > 10$", is true, it follows that $\exists x P(x)$ is true.■

●Translating from Formal to Informal Language

Example.

Rewrite the following formal statements in a variety of equivalent but more informal ways. Do not use the symbol ∀ or ∃.

(a) $\forall x \in \mathbb{R}, x^2 \geq 0$;

(b) $\forall x \in \mathbb{R}, x^2 \neq -1$;

(c) $\exists m \in \mathbb{Z}$ such that $m^2 = m$.

Solution.

(a) Every real number has a nonnegative square.

Or: All real numbers have nonnegative squares.

Or: Any real number has a nonnegative square.

Or: The square of each real number is nonnegative.

(b) All real numbers have squares that do not equal -1.

Or: No real numbers have squares equal to -1.

(The words none are or no … are equivalent to the words all are not.)

(c) There is a positive integer whose square is equal to itself.

Or: We can find at least one positive integer equal to its own square.

Or: Some positive integer equals its own square.

Or: Some positive integers equal their own squares. ■

Another way to restate universal and existential statements informally is to place the quantification at the end of the sentence. For instance, instead of saying "For any real number *x*, x^2 is nonnegative," you could say "x^2 is nonnegative for any real number *x*." In such a case the quantifier is said to "**trail**" the rest of the sentence.

●Trailing Quantifiers

Example.

Rewrite the following statements so that the quantifier trails the rest of the sentence.

(a) For any integer n, $2n$ is even.

(b) There exists at least one real number x such that $x^2 \leq 0$.

Solution.

(a) $2n$ is even for any integer n.

(b) $x^2 \leq 0$ for some real number *x*.

Or: $x^2 \leq 0$ for at least one real number *x*. ■

●Translating from Informal to Formal Language

Example.

Rewrite each of the following statements formally. Use quantifiers and variables.

(a) All triangles have three sides.

(b) No dogs have wings.

(c) Some programs are structured.

Solution.

(a) ∀ triangle *t*, *t* has three sides.

Or: $\forall t \in T$, *t* has three sides (where *T* is the set of all triangles).

(b) ∀ dog *d*, *d* does not have wings.

Or: $\forall d \in D$, *d* does not have wings (where *D* is the set of all dogs).

(c) ∃ a program *p* such that *p* is structured.

Or: $\exists p \in P$ such that *p* is structured (where *P* is the set of all programs). ■

●Universal Conditional Statements

A reasonable argument can be made that the most important form of statement in mathematics is the **universal conditional statement**:

$$\forall x, \text{ if } P(x) \text{ then } Q(x).$$

Familiarity with statements of this form is essential if you are to learn to speak mathematics.

●Writing Universal Conditional Statements Informally

Example.

Rewrite the following statement informally, without quantifiers or variables.

$\forall x \in \mathbb{R}$, if $x > 2$, then $x^2 > 4$.

Solution.

If a real number is greater than 2, then its square is greater than 4.

Or: Whenever a real number is greater than 2, its square is greater than 4.

Or: The square of any real number greater than 2 is greater than 4.

Or: The squares of all real numbers greater than 2 are greater than 4. ■

Example.

Rewrite each of the following statements in the form

∀ … …., if……., then……… .

(a) If a real number is an integer, then it is a rational number.

(b) All bytes have eight bits.

(c) No fire trucks are green.

Solution.

(a) ∀ real number *x*, if *x* is an integer, then *x* is a rational number.

Or: $\forall x \in \mathbb{R}$, if $x \in \mathbb{Z}$ then $x \in \mathbb{Q}$.

(b) $\forall x$, if *x* is a byte, then *x* has eight bits.

(c) $\forall x$, if *x* is a fire truck, then *x* is not green. ■

●Equivalent Forms of Universal and Existential Statements

Observe that the two statements

“∀ real number *x*, if *x* is an integer then *x* is rational”

and

“∀ integer *x*, *x* is rational”

mean the same thing because the set of integers is a subset of the set of real numbers. Both have informal translations

“All integers are rational.”

In fact, a statement of the form

$$\forall x \in U, \text{ if } P(x) \text{ then } Q(x)$$

can always be rewritten in the form

$$\forall x \in D, Q(x)$$

by narrowing *U* to be the subset *D* consisting of all values of the variable *x* that make $P(x)$ true. Conversely, a statement of the form

$$\forall x \in D, Q(x)$$

can be rewritten a

$$\forall x, \text{ if } x \text{ is in } D \text{ then } Q(x)$$

Example.

Rewrite the following statement in the two forms

“$\forall x$, if…….. then………. ”

and

“$\forall \dots \dots \dots x$,………. ”:

“All squares are rectangles” .

Solution.

“$\forall x$, if *x* is a square then *x* is a rectangle”.

and

“$\forall$ square x, *x* is a rectangle”. ■

Similarly, a statement of the form

“$\exists x$ such that $P(x)$ and $Q(x)$”

can be rewritten as

“$\exists x \in D$ such that $Q(x)$,”

where *D* is the set of all *x* for which $P(x)$ is true.

Example.

A **prime number** is an integer greater than 1 whose only positive integer factors are itself and 1.

Consider the statement

“There is an integer that is both prime and even.”

Let $P(n)$ be “*n* is prime” and $E(n)$ be “*n* is even.”

Use the notation $P(n)$ and $E(n)$ to rewrite this statement in the following two forms:

a. $\exists n$ such that … … … $\wedge$ … … ….

b. $\exists$ … … … n such that ……….

Solution.

(a) $\exists n$ such that $P(n) \wedge E(n)$.

(b) Two answers:

$\exists$ a prime number *n* such that $E(n)$.

$\exists$ an even number *n* such that $P(n)$. ■

Example.

What do the following statements mean, where the domain in each case consists of the real numbers?

(1) $\forall x < 0(x^2 > 0)$;

(2) $\forall y \neq 0(y^3 \neq 0)$;

(3) and $\exists z > 0(z^2 = 2)$.

Solution.

(1) The statement $\forall x < 0(x^2 > 0)$ states that for every real number *x* with $x < 0$, $x^2 > 0$. That is, it states "The square of a negative real number is positive".

This statement is the same as $\forall x\big(x < 0 \rightarrow (x^2 > 0)\big)$.

(2) The statement $\forall y \neq 0\ (y^3 \neq 0)$, states that for every real number y with $y \neq 0$, we have $y^3 \neq 0$ that is, it states

"the cube of every nonzero real number is nonzero."

Note that this statement is equivalent to

$$\forall y(y \neq 0 \longrightarrow y^3 \neq 0).$$

(3) The statement $\exists z > 0(z^2 = 2)$ states that there exists a real number z with $z > 0$ such that $z^2 = 2$. That is, it states

"there is a positive root of 2."

This statement is equivalent to $\exists z(z > 0 \wedge z^2 = 2)$.■

● Precedence of Quantifiers

The quantifiers ∀ and ∃ have higher precedence than all logical operators from propositional calculus. For example, $\forall x P(x) \vee Q(x)$ is the disjunction of $\forall x P(x)$ and $Q(x)$. In other words, it means $(\forall x P(x)) \vee Q(x)$ rather than $\forall x(P(x) \vee Q(x))$.

♣Logical Equivalence Involving Quantifiers

Definition.

Statements involving predicates and quantifiers are **logically equivalent** if and only if they have the same truth value no matter which predicates are substituted into these statements. We use the notation $S \equiv T$ to indicate that two statements S and T involving predicates and quantifiers are logically equivalent.

Example.

Show that $\forall x(P(x) \wedge Q(x))$ and $\forall x P(x) \wedge \forall x Q(x)$ are logically equivalent, where the same domain is used throughout.

Solution.

To show that these statements are logically equivalent, we must show that they always take the same truth value, no matter what predicate P and Q are, and no matter which domain of discourse is used.

Suppose we have particular predicates P and Q, with a common domain. We can show that $\forall x(P(x) \wedge Q(x))$ and $\forall x P(x) \wedge \forall x Q(x)$ are logically equivalent by doing two things. First, we show that if $\forall x(P(x) \wedge Q(x))$ is true, then $\forall xP(x) \wedge \forall xQ(x)$ is true.

Second, we show that if $\forall xP(x) \wedge \forall xQ(x)$ is true, then $\forall x(P(x) \wedge Q(x))$ is true.

So, suppose that $\forall x(P(x) \wedge Q(x))$ is true. This means that if a is in the domain, then $P(a) \wedge Q(a)$ is true. Hence $P(a)$ is true and $Q(a)$. Because $P(a)$ is true and $Q(a)$ for every element in the domain, we can conclude that $\forall xP(x)$ and $\forall xQ(x)$ are both true. This means that $\forall xP(x) \wedge \forall xQ(x)$ is true.

Next, suppose that $\forall xP(x) \wedge \forall xQ(x)$ is true. It follows that $\forall xP(x)$ is true and $\forall xQ(x)$ is true. Hence if a is in the domain, then $P(a)$ is true and $Q(a)$ is true. It follows that for all a, $P(a) \wedge Q(a)$ is true. It follows that $\forall x(P(x) \wedge Q(x))$ is true.

Therefore $\forall x\big(P(x) \wedge Q(x)\big) \equiv \forall xP(x) \wedge \forall xQ(x)$. ■

Exercise.

Prove that $\exists x\big(p(x) \vee Q(x)\big) \equiv \exists xp(x) \vee \exists xQ(x)$, where the same domain is used throughout.

● Negating Quantifier Expressions

We will often want to consider the negation of a quantified expression. For instance, consider the negation of the statement

"Every student in your class has taken a course in calculus"

This statement is a universal quantification, namely $\forall x P(x)$, where $P(x)$ is the statement

"x has taken a course in calculus"

and the domain consists of the students in your class. The negation of this statement is

"It is not the case that every student in your class has taken a course in calculus".

This is equivalent to

"There is a student in your class who has not taken a course in calculus".

And this is simply the existential quantification of the negation of the original propositional function, namely, $\exists x \neg P(x)$. This example illustrates the following equivalence

$$\neg \forall x P(x) \equiv \exists x \neg P(x).$$

Example.

Prove that:

$$\neg\forall x P(x) \equiv \exists x \neg P(x).$$

Where the same domain is used throughout.

Proof.

To show that $\neg\forall x P(x)$ and $\exists x \neg P(x)$ are logically equivalent no matter what the propositional function $P(x)$ is and what the domain is.

First note that $\neg\forall x P(x)$ is true if and only if $\forall x P(x)$ is false.

Next, note that $\forall x P(x)$ is false if and only if there is an element x in the domain for which $P(x)$ is false.

This holds if and only if there is an element x in the domain for which $\neg P(x)$ is true.

Finally, note that there is an element x in the domain for which $\neg P(x)$ is true if and only if $\exists x \neg P(x)$ is true.

It follows that $\neg\forall x P(x)$ and $\exists x \neg P(x)$ are logically equivalent. ■

Suppose we wish to negate an existential quantification. For instance, consider the statement

“There is a student in this class who has taken a course in calculus”.

This is the existential quantification $\exists x Q(x)$ where $Q(x)$ is the statement

“x has taken a course in calculus”.

The negation of this statement is

“It is not the case that there is a student in this class who has taken calculus”

which is just the universal quantification of the negation of the original propositional function, or, $\forall x \neg Q(x)$.

This example illustrates the equivalence:

$$\neg \exists x Q(x) \equiv \forall x \neg Q(x).$$

Exercise.

Prove that:

$$\neg \exists x Q(x) \equiv \forall x \neg Q(x),$$

where the same domain is used throughout.

Example.

What is the negation of the following statements

(a) $\forall x(x^2 > x)$;

(b) $\exists x(x^2 = 2)$.

Solution.

(a) The negation of $\forall x(x^2 > x)$ is the statement $\neg\forall x(x^2 > x)$, which is equivalent to $\exists x\neg(x^2 > x)$.

This can be rewritten as $\exists x(x^2 \leq x)$.

(b) The negation of $\exists x(x^2 = 2)$ is the statement $\neg\exists x(x^2 = 2)$, which is equivalent to $\forall x\neg(x^2 = 2)$.

This can be rewritten as $\forall x(x^2 \neq 2)$.

The truth values of the statements in (1) and (2) depend on the domain of discourse. ■

♣ Now, we give some examples to show how to translate sentences into logical expressions.

Example.

Express the statement

“Every student in this class has studied calculus”

using predicates and quantifiers.

Solution.

First, we rewrite the statement so that we can clearly identify the appropriate quantifiers to use. Doing so, we obtain:

“For every student in this class, that student has studied calculus”.

Next, we introduce a variable x so that our statement becomes

“for every student x in this class, x has studied calculus”.

Continuing, we introduce the predicate $C(x)$, which is the statement

“x has studied calculus”.

Consequently, if the domain of discourse for x consists of the students in the class, we can translate our statement as

$$\forall x C(x).$$ ■

Example.

Express the statement

"Some student in this class has visited Cairo",

and

"Every student in this class has visited either Cairo or Alexandria".

Solution.

The statement "Some student in this class has visited Cairo" means that "There is a student in this class with the property that the student has visited Cairo".

We can introduce a variable x, so that our statement becomes "There is a student x in this class having the property x has visited Cairo". We introduce the predicate $M(x)$, which is the statement "x has visited Cairo". If the domain of discourse of x consists of the students in this class, we can translate this first statement as $\exists x M(x)$. Similarly, the second statement can be expressed as $\forall x \big(C(x) \vee M(x)\big)$, where the domain of discourse of x consists of all students in this class, $M(x)$ be the statement "x visited Cairo" and $C(x)$ be the statement "x visited Alexandria". ■

Example.

Write formal negations for the following statements:

(a) ∀ primes *p*, *p* is odd;

(b) ∃ a triangle *T* such that the sum of the angles of *T* equals 200°.

Solution.

(a) By applying the rule for the negation of a ∀ statement, you can see that the answer is ∃ a prime *p* such that *p* is not odd.

(b) By applying the rule for the negation of a ∃ statement, you can see that the answer is ∀ triangles *T*, the sum of the angles of *T* does not equal 200°. ■

Example.

Rewrite the following statements formally. Then write formal and informal negations.

(a) No politicians are honest;

(b) The number 1357 is not divisible by any integer between 1 and 37.

Solution

(a) *Formal version*: ∀ politicians *x*, *x* is not honest.

Formal negation: ∃ a politician *x* such that *x* is honest.

Informal negation: Some politicians are honest.

(b) This statement has a trailing quantifier.

Written formally it becomes:

∀ integer *n* between 1 and 37, 1357 is not divisible by *n*.

Its negation is therefore

∃ an integer *n* between 1 and 37 such that 1357 is divisible by *n*.

An informal version of the negation is

The number 1,357 is divisible by some integer between 1 and 37. ■

Example.

Write informal negations for the following statements:

a. All computer programs are finite.

b. Some computer hackers are over 40.

Solution.

a. What exactly would it mean for this statement to be false? The statement asserts that all computer programs satisfy a certain property. So for it to be false, there would have to be at least one computer program that does not satisfy the property. Thus the answer is

There is a computer program that is not finite.

Or: Some computer programs are infinite.

b. This statement is equivalent to saying that there is at least one computer hacker with a certain property. So for it to be false, not a single computer hacker can have that property. Thus the negation is

No computer hackers are over 40.

Or: All computer hackers are 40 or under. ■

●Negations of Universal Conditional Statements

Negations of universal conditional statements are of special importance in mathematics. The form of such negations can be derived from facts that have already been established. By definition of the negation of a *for all* statement, $\neg\forall x(P(x) \to Q(x)) \equiv \exists x\neg(P(x) \to Q(x))$; But the negation of an if-then statement is logically equivalent to an and statement. More precisely,

$$\neg(P(x) \to Q(x)) \equiv P(x) \wedge \neg Q(x).$$

Therefore, $\neg\forall x(P(x) \to Q(x)) \equiv \exists x(P(x) \wedge \neg Q(x))$.

Example.

Write a formal negation for statement (a) and an informal negation for statement (b).

a. ∀ person *p*, if *p* is blond then *p* has blue eyes.

b. If a computer program has more than 100,000 lines, then it contains a bug.

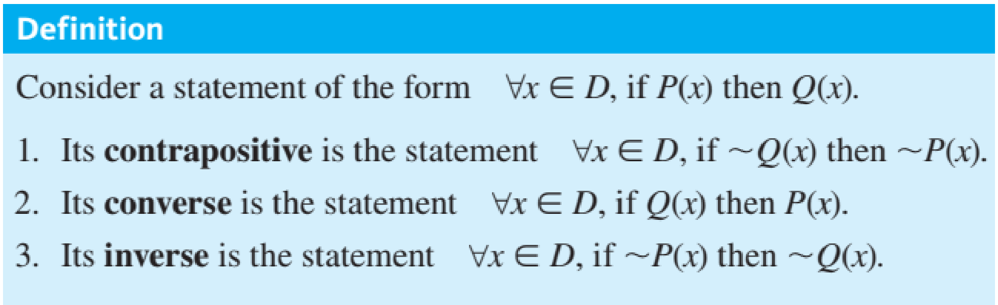
Solution.

a. ∃ a person *p* such that *p* is blond and *p* does not have blue eyes.

b. There is at least one computer program that has more than 100,000 lines and does not contain a bug.■

Variants of Universal Conditional Statements

Recall that a conditional statement has a contrapositive, a converse, and an inverse. The definitions of these terms can be extended to universal conditional statements.



Definition

Consider a statement of the form $\forall x \in D$, if $P(x)$ then $Q(x)$.

1. Its **contrapositive** is the statement $\forall x \in D$, if $\sim Q(x)$ then $\sim P(x)$.
2. Its **converse** is the statement $\forall x \in D$, if $Q(x)$ then $P(x)$.
3. Its **inverse** is the statement $\forall x \in D$, if $\sim P(x)$ then $\sim Q(x)$.

Example.

Write a formal and an informal contrapositive, converse, and inverse for the following statement:

If a real number is greater than 2, then its square is greater than 4.

Solution.

The formal version of this statement is:

$$\forall x \in \mathbb{R}(x > 2 \rightarrow x^2 > 4).$$

Contrapositive: $\forall x \in \mathbb{R}(x^2 \leq 4 \rightarrow x \leq 2)$.

Or: If the square of a real number is less than or equal to 4, then the number is less than or equal to 2.

Converse: $\forall x \in \mathbb{R}(x^2 > 4 \rightarrow x > 2)$.

Or: If the square of a real number is greater than 4, then the number is greater than 2.

Inverse: $\forall x \in \mathbb{R}(x \leq 2 \rightarrow x^2 \leq 4)$..

Or: If a real number is less than or equal to 2, then the square of the number is less than or equal to 4.

Note that in solving this example, we have used the equivalence of "$x \ngtr a$" and "$x \leq a$" for all real numbers x and a. ■

Exercise.

(a) Prove that a universal conditional statement is logically equivalent to its contrapositive.

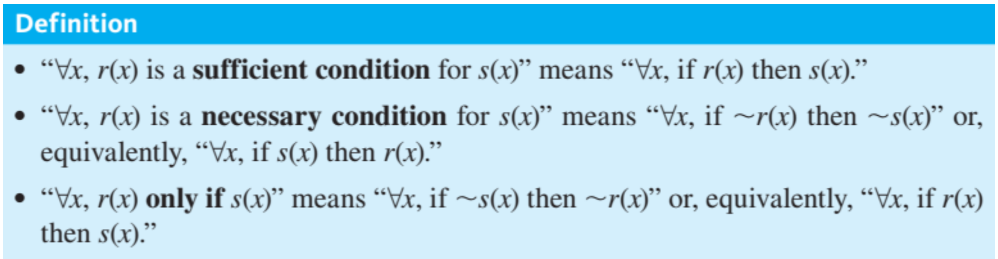
(b) Prove that a universal conditional statement is not logically equivalent to its converse.

(c) Prove that a universal conditional statement is not logically equivalent to its inverse.

Note that answering of both (b) and (c) is by giving counterexamples.

●Necessary and Sufficient Conditions, Only If

The definitions of *necessary*, *sufficient*, and *only if* can also be extended to apply to universal conditional statements.



Definition

- "$\forall x$, $r(x)$ is a **sufficient condition** for $s(x)$" means "$\forall x$, if $r(x)$ then $s(x)$."
- "$\forall x$, $r(x)$ is a **necessary condition** for $s(x)$" means "$\forall x$, if $\sim r(x)$ then $\sim s(x)$" or, equivalently, "$\forall x$, if $s(x)$ then $r(x)$."
- "$\forall x$, $r(x)$ **only if** $s(x)$" means "$\forall x$, if $\sim s(x)$ then $\sim r(x)$" or, equivalently, "$\forall x$, if $r(x)$ then $s(x)$."

Example.

Rewrite each of the following as a universal conditional statement, quantified either explicitly or implicitly. Do not use the word *necessary* or *sufficient*.

a. Squareness is a sufficient condition for rectangularity.

b. Being at least 35 years old is a necessary condition for being president of the Egypt.

Solution.

a. A formal version of the statement is:

$\forall x$, if x is a square, then x is a rectangle.

Or, with implicit universal quantification:

If a figure is a square, then it is a rectangle.

b. Using formal language, you could write the answer as

∀ person *x*, if *x* is younger than 35, then *x* cannot be president of the Egypt.

Or, by the equivalence between a statement and its contrapositive:

∀ person *x*, if *x* is president of the United States, then *x* is at least 35 years old. ■

Example.

Rewrite the following as a universal conditional statement:

A product of two numbers is 0 only if one of the numbers is 0.

Solution.

Using informal language, you could write the answer as If it is not the case that one of two numbers is 0, then the product of the numbers is not 0. In other words, If neither of two numbers is 0, then the product of the numbers is not 0. Or, by the equivalence between a statement and its contrapositive. If a product of two numbers is 0, then one of the numbers is 0. ■

♣ Nested Quantifiers

In this section, we will study **nested quantifiers**, which are quantifiers that occur within the scope of other quantifiers, such as in the statement

$$\forall x \exists y (x + y = 0).$$

Nested quantifiers commonly occur in mathematics and computer science.

Example.

Assume that the domain for the variables x and y consists of all real numbers. The statement

$$\forall x \forall y (x + y = y + x)$$

says that $x + y = y + x$ for all real numbers x and y. This is the commutative law for addition of real numbers. Likewise, the statement $\forall x \exists y (x + y = 0)$ says that for every real number x there is a real number y such that $x + y = 0$. This states that every real number has an additive inverse. The statement

$$\forall x \forall y \forall z [x + (y + z) = (x + y) + z]$$

is the associative law for addition of real numbers. ■

Example.

Translate into English **(Informal Language)** the statement

$$\forall x \forall y (x > 0) \wedge (y < 0) \longrightarrow (xy < 0),$$

where domain for both variables consists of all real numbers.

Solution.

This statement says that for every real number x and for every real number y, if $x > 0$ and $y < 0$, then $xy < 0$. That is, this statement says that for real numbers x and y, if x is positive and y is negative, then x y is negative. This can be stated more succinctly as

"The product of a positive real number and a negative real number is a negative real number". ■

Example.

The **reciprocal** of a real number a is a real number b such that $ab = 1$. The following two statements are true. Rewrite them **formally** using quantifiers and variables.

a. Every nonzero real number has a reciprocal.

b. There is a real number with no reciprocal.

Solution.

a. ∀ nonzero real number u, ∃ a real number v such that $uv = 1$.

Equivalently,

$$\forall u \exists v (uv = 1),$$

where domain for both variables consists of all real numbers.

b. ∃ a real number c such that ∀ real number $d, cd \neq 1$.

Equivalently,

$$\exists u \forall v (uv \neq 1),$$

where domain for both variables consists of all real numbers. ■

Example.

Consider the statement

"There is a smallest positive integer."

Write this statement formally using both symbols ∃ and ∀.

Solution.

To say that there is a smallest positive integer means that there is a positive integer m with the property that no matter what positive integer n a person might pick, m

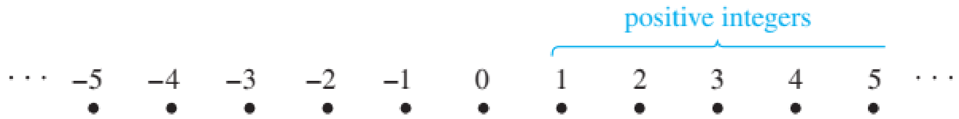
will be less than or equal to n:

$\exists$ a positive integer m such that $\forall$ positive integer n, $m \leq n$.

Equivalently,

$$\exists m \forall n (m \leq n), \text{ where } m, n \in \mathbb{Z}^+.$$

Note that this statement is true because 1 is a positive integer that is less than or equal to every positive integer.



■

Example.

Consider the statement

“There is no smallest positive real number.”

Write this statement formally using both symbols $\exists$ and $\forall$.

Solution.

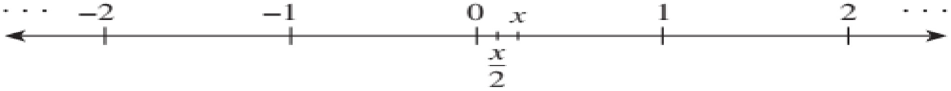
$\forall$ positive real number x, $\exists$ positive real number y, such that $y < x$.

Equivalently,

$$\forall x \exists y (y < x), \text{ where } x, y \in \mathbb{R}^+.$$

Note that this statement is true.

Imagine the positive real numbers on the real number line. These numbers correspond to all the points to the right of 0. Observe that no matter how small a real number x is, the number $x/2$ will be both positive and less than x.



■

● The Order of Quantifiers

Many mathematical statements involve multiple quantifications of propositional functions involving more than one variable. It is important to note that the order of the quantifiers is important, unless all the quantifiers are universal quantifiers or all are existential quantifiers.

Example.

Consider the following two statements:

∀ person *x*, ∃ a person *y* such that *x* loves *y*.

∃ a person *y* such that ∀ person *x*, *x* loves *y*.

Note that except for the order of the quantifiers, these statements are identical. However, the first means that given any person, it is possible to find someone whom that person loves, whereas the second means that there is one amazing individual who is loved by all people. (Reread the statements carefully to verify these interpretations!) The two sentences illustrate an extremely important property about statements with two different quantifiers. ■

Example.

Consider the commutative property of addition of real numbers, for example:

∀ real number x and ∀ real number y, $x + y = y + x$.

$$\forall x \forall y\ (x + y = y + x).$$

This means the same as

∀ real number y and ∀ real number x, $x + y = y + x$.

$$\forall x \forall y\ (x + y = y + x).$$ ■

Example.

Translate the statement

$$\forall x (C(x) \lor \exists y \big(C(y) \land F(x, y)\big))$$

into English, where $C(x)$ is "x has a computer", $F(x, y)$ is "x and y are friends" and the domain for both x and y consists of all students in your faculty.

Solution.

The statement says that for every student *x* in your faculty *x* has a computer or there is a student *y* such that *y* has a computer and *x* and *y* are friends. In other words, every student in your faculty has a computer or has a friend who has a computer. ■

Example.

Express the statement "If a person is female and is a parent, then this person is someone's mother" as a logical expression involving predicates, quantifiers with a domain consisting of all people.

Solution.

The statement "If a person is female and is a parent, then this person is someone's mother" can be expressed as "For every person *x*, if person *x* is female and person *x* is a parent, then there exists a person *y* such that person *x* is the mother of person *y*". We introduce the predicates $F(x)$ to represent "*x* is female" $P(x)$ to represent "*x* is a parent" and $M(x, y)$ to represent "*x* is the mother of *y*". The original statement can be represent by

$$\forall x\big(F(x) \wedge P(x) \longrightarrow \exists y M(x,y)\big)$$

We can move $\exists y$ all the way to the left, because y does not appear in $F(x) \wedge P(x)$, to obtain an equivalent expression $\forall x \exists y\big(F(x) \wedge P(x) \longrightarrow M(x,y)\big)$.■

Example.

Express the statement "Everyone has exactly one best friend" as a logical expression involving predicates, quantifiers with a universe of discourse consisting of all people and logical connectives.

Solution.

The given statement can be expressed as "For every person *x*, person *x* has exactly one best friend". Introducing the universal quantifier, we see that this statement is the same as "$\forall x$(person *x* has exactly one best friend)" where the universe of discourse consists of all people. To say that *x* has exactly one best friend means that there is a person *y* who is the best friend of *x*, and furthermore, that for every person *z*, if person *z* is not person *y*, then *z* is not the best friends of *x*. When we introduce the predicate $B(x,y)$ to be the statement "*y* is the best friend of *x*” the statement *x* has exactly best friend can be represented as

$$\exists y\left((B(x,y) \wedge \forall z\big((z \neq y) \longrightarrow \neg B(x,z)\big)\right)$$

Consequently, our original statement can be expressed as

$$\forall x \exists y\left(B(x,y) \wedge \forall z\big((z \neq y) \longrightarrow \neg B(x,z)\big)\right).$$ ■

Example.

Let $P(x,y)$ be the statement “$x + y = y + x$”. What is the truth value of the quantifications $\forall x \forall y P(x,y)$ and $\forall y \forall x P(x,y)$, where the domain for all variables consists of all real numbers?

Solution.

The quantification $\forall x \forall y P(x,y)$ denotes the proposition “for all real numbers *x* and for all real numbers *y*, $x + y = y + x$". Since $P(x,y)$ is true for all real numbers *x* and *y*, the proposition $\forall x \forall y P(x,y)$ is true. Note that $\forall y \forall x P(x,y)$ says "For all real numbers *y*, for all real numbers *x*, $x + y = y + x$”. This has the same meaning as the statement as “For all real numbers *x* and for all real numbers $y, x + y = y + x$”. That is, $\forall x \forall y P(x,y)$

and $\forall y \forall x P(x, y)$ have the same meaning, and both are true. ■

Example.

Let $Q(x, y)$ denote "$x + y = 0$".What are the truth value of the quantifications $\exists y \forall x Q(x, y)$ and $\forall x \exists y Q(x, y)$, where the domain for all variables consists of all real numbers?

Solution.

The quantification $\exists y \forall x Q(x, y)$ denotes the proposition "there is a real number *y* such that for every real number $x, Q(x, y)$". No matter what value of *y* is chosen, there is only one value of *x* for which $x + y = 0$. Since there is no real number y such that $x + y = 0$ for all real numbers x, the statement $\exists y \forall x Q(x, y)$ is false.

The quantification $\forall x \exists y Q(x, y)$ denotes the proposition "for every real number *x* there is a real number *y* such that $x + y = 0$, namely $y = -x$. Hence the statement $\forall x \exists y Q(x, y)$ is true. ■

Note. The above example illustrates that the statements $\exists y \forall x Q(x, y)$ and $\forall x \exists y Q(x, y)$ are not logically equivalent.

●Negating Nested Quantifiers

Statements involving nested quantifiers can be negated by successively applying the rules for negating statements involving a single quantifier.

Example.

Write a negation for each of the following statements, and determine which is true, the given statement or its negation.

(a) For every square *x*, there is a circle *y* such that *x* and *y* have the same color.

(b) There is a triangle *x* such that for every square *y*, *x* is to the right of *y*.

Solution.

(a) *First version of negation*:

∃ a square *x* such that , ¬ (∃ a circle *y* such that *x* and *y* have the same color).

Final version of negation:

∃ a square *x* such that ∀ circle *y*, *x* and *y* do not have the same color.

(b) *First version of negation*:

∀ triangle x, ¬ (∀ square y, x is to the right of y).

Final version of negation:

∀ triangle x, ∃ a square y such that x is not to the right of y. ■

Example.

Express the negation of the statement $\forall x\exists y(xy = 1)$ so that no negation precedes a quantifier.

Solution.

$$\neg\forall x\exists y(xy = 1) \equiv \exists x\neg\big(\exists y(xy = 1)\big)$$

$$\equiv \exists x\forall y\neg(xy = 1) \equiv \exists x\forall y(xy \neq 1).$$ ■

The following table summarizes the meanings of the different possible quantifications involving two variables.

Statement	When True?	When False?
$\forall x\forall yP(x,y)$ $\forall y\forall xP(x,y)$	$P(x,y)$ is true for every pair x,y	There is a pair x,y for which $P(x,y)$ is false
$\forall x\exists yP(x,y)$	For every x there is a y for which $P(x,y)$ is true	There is an x such that $P(x,y)$ is false for every y.
$\exists x\forall yP(x,y)$	There is an x for which $P(x,y)$ is true for every y.	For every x there is a y for which $P(x,y)$ is false.
$\exists x\exists yP(x,y)$ $\exists y\exists xP(x,y)$	There is a pair x,y for which $P(x,y)$ is true.	$P(x,y)$ is false for every pair x,y.

♣Solved problems

1. Determine whether the following proposition are true or false:

(a) $\forall x \in D(2^x < x!)$, $D = \{1,2,3\}$; (false :$x = 1$)

(b) $\exists x \in D(2^x < x!)$, $D = \{1,2,3,4,5\}$; (true : $x = 1$)

(c) $\forall x \in D(2^x < x!)$, $D = \{4,5,6\}$; (true :)

(d) $\forall x \in D\left(x > \frac{1}{x}\right)$, $D = \mathbb{R}^*$ (false : $x > \frac{1}{x}$)

(e) $\exists x \in D\left(x > \frac{1}{x}\right)$, $D = \mathbb{R}$ (true: $x = 2$)

(f) $\forall x \in D(x^2 \neq x + 2)$, $D = \mathbb{Q}$ (false: $x = 2$)

(g) $\exists x \in D(x^2 = 2)$, $D = \mathbb{Q}$ (false: $x = \pm\sqrt{2}$)

(h) $\forall x \in D(x^2 + X + 41 \text{ is prime})$ (false :$x = 41$)

(k) $\exists x \in D(x^2 + x + 1 = 0)$, $D = \mathbb{R}$ (false: $x = \frac{-1\pm\sqrt{3}}{2}$)

2. Prove that:

(a) $\neg\big(\exists x P(x)\big) \equiv \forall x\big(\neg P(x)\big)$,

$\neg\big(\forall x P(x)\big) \equiv \exists x \neg P(x)$

(b) $\neg\big(\forall x P(x) \longrightarrow Q(x)\big) \equiv \exists x\big(P(x) \wedge \neg Q(x)\big)$,

$\neg(\exists x(P(x) \wedge Q(x))) \equiv \forall x\big(P(x) \longrightarrow \neg Q(x)\big)$

Solution.

(b) $\neg\big[\forall x\big(P(x) \longrightarrow Q(x)\big)\big] \equiv \neg\forall x\big(\neg P(x) \vee Q(x)\big)$

$$\equiv \exists x \neg\big(\neg P(x) \vee Q(x)\big) \equiv \exists x\big(\neg\neg P(x) \wedge \neg Q(x)\big)$$

$$\equiv \exists x\big(P(x) \wedge \neg Q(x)\big)\ . ■$$

3. Negate the following statements:

(a) There is no one in the island;

(b) Every real number x satisfies $x^2 = 1$;

(c) Students who likes Mathematics likes physics, too;

(d) All student and staff came to the meeting.

Solution.

(a) Let $M(x)$ is "x is one", $I(x)$ is "x is in the island" so , the given statement is $\forall x\big(M(x) \longrightarrow \neg I(x)\big)$.

Therefore

$$\forall x\big(M(x) \longrightarrow \neg I(x)\big) \equiv \forall x\big(\neg M(x) \vee \neg I(x)\big)$$

$$\equiv \forall x \neg\big(M(x) \wedge I(x)\big)$$

$$\equiv \neg\exists x\big(M(x) \wedge I(x)\big)$$

Then, the negation of the given statement is

$\exists x\big(M(x) \wedge I(x)\big)$,

or "There is someone in the island".

(b) The given statement is $\forall x(R(x) \longrightarrow x^2 = 1)$, where $R(x)$: is x real.

So, its negation is

$$\neg\forall x(R(x) \longrightarrow x^2 = 1) \equiv \neg\forall x(\neg R(x) \vee x^2 = 1)$$
$$\equiv \exists x(R(x) \wedge x^2 \neq 1).$$

Therefore, the negation of our statement is

"There exists a real number *x* such that $x^2 \neq 1$"

(c) "Students who likes Math. likes Phys., too " is

$\forall x\big(M(x) \longrightarrow P(x)\big)$, where x is student, $M(x)$: likes Math., $P(x)$: likes phys. Therefore

$$\neg\forall x\big(M(x) \longrightarrow P(x)\big) \equiv \exists x\big(M(x) \wedge \neg P(x)\big)$$

or

"Some students like Mathematics but not like Physics”

(d) $S(x)$: x is student; $T(x)$: x is a teacher

$M(x)$: x came to the meeting

The given statement is

$$\forall x\Big[\big(S(x) \longrightarrow M(x)\big) \wedge \big(T(x) \longrightarrow M(x)\big)\Big] .$$

Therefore

$$\neg\forall x\Big[\big(S(x) \longrightarrow M(x)\big) \wedge \big(T(x) \longrightarrow M(x)\big)\Big]$$
$$\equiv \exists x\neg\Big[\big(S(x) \longrightarrow M(x)\big) \wedge \big(T(x) \longrightarrow M(x)\big)\Big]$$
$$\equiv \exists x\neg\Big[\big(\neg S(x) \vee M(x)\big) \wedge \big(\neg T(x) \vee M(x)\big)\Big]$$
$$\equiv \exists x[\neg(\neg S(x) \vee M(x)\ ) \vee \neg(\neg T(x) \vee M(x)\ )\ ]$$
$$\equiv \exists x\Big[\big(S(x) \wedge \neg M(x)\big) \vee \big(T(x) \wedge \neg M(x)\big)\Big]$$

$$\equiv \exists x\left(\left(S(x)\vee T(x)\right)\wedge\neg M(x)\right)$$

Or,

"Some students or some staffs do not come to the meeting". ■

4. Translate each of these statements into logical expressions using predicates, quantifiers and logical connectives.

(a) Every real number is complete square.

(b) There is a real number between each pair of distinct real numbers.

(c) There is a multiple integer for the number 5 but not a multiple for 7.

Solution.

(a) $(\forall x)(\exists y)(x = y^2)$. The domain forx, y is $\mathbb{R}$.

(b) $(\forall x)(\forall y)\left(x \neq y \longrightarrow \exists z\left((x < z \wedge z < y) \vee (y < z \wedge z < x)\right)\right)$

(c) $\exists x(\exists y(x = 5y) \wedge \forall z(x \neq 7z)\ )$,

where the domain of x, y, z is $\mathbb{Z}$. ■

5. Translate the following statements into ordinary language, where the domain for x, y is $\mathbb{R}$.

(a) $(\exists x)(\forall y)(y < x)$;

(b) $(\forall x)(\exists y)(x < y)$.

Solution.

(a)There is a real number greater than all real.

(b) For each real number, there is another real number greater than it.

6. Express the following statement using logical operators, predicate and quantifiers.

(a) "For every real $\mathcal{E} > 0$, We can find integer *k* such that : if $n > k$, then a_n lies between $L + \mathcal{E}$ and $L - \mathcal{E}$.

(b) Negate the statement in (a).

Solution.

$$\forall \mathcal{E}(\mathcal{E} > 0 \rightarrow (\exists k)(\forall n)\big(n > k \\ \rightarrow (L - \mathcal{E} < a_n) \wedge (a_n < L + \mathcal{E})\big))$$

The domain of $\mathcal{E}$ is $\mathbb{R}$ and the domain of *k*, *n* is $\mathbb{Z}$.

The negation:

$$\neg\forall \mathcal{E}(\mathcal{E} > 0 \rightarrow (\exists k)(\forall n)(n > k \\ \rightarrow (L - \mathcal{E} < a_n) \wedge (a_n < L + \mathcal{E})))$$

$$\equiv \neg\forall \mathcal{E}(\neg(\mathcal{E} > 0) \vee (\exists k)(\forall n)(n > k \\ \rightarrow (L - \mathcal{E} < a_n) \wedge (a_n < L + \ \mathcal{E})))$$

$$\equiv \exists \mathcal{E}(\mathcal{E} > 0 \land \neg(\exists k)(\forall n)(n > k$$
$$\to (L - \mathcal{E} < a_n) \land (a_n < L + \ \mathcal{E})))$$

$$\equiv \exists \mathcal{E}(\mathcal{E} > 0 \land (\forall k)(\exists n)\neg(n > k$$
$$\to (L - \mathcal{E} < a_n) \land (a_n < L + \ \mathcal{E})))$$

$$\equiv \exists \mathcal{E}(\mathcal{E} > 0 \land (\forall k)(\exists n)(n$$
$$> k \land (L - \mathcal{E} \geq a_n) \lor (a_n \geq L + \ \mathcal{E})))$$

There is $\varepsilon > 0$ such that for every integer k there exist $n > k$ with either $L - \varepsilon \geq a_n$ or $a_n \geq L + \ \varepsilon$. ■

6. Determine whether the following proposition are true or false:

(a) $Q(x, y) = (\forall x)(\exists y)(x \leq y), D_x = D_y = \mathbb{R}$.

(b) $P(x, y) = (\forall x)(\forall y)\big(x \leq y \longrightarrow \neg(x \leq y)\big),$

$$D_x = D_y = \mathbb{R}$$

(c) $R(x, y) = (\exists x)(\exists y)(x + 5 = y^2),\ D_x = D_y = \mathbb{R}$.

(d) $Z(x, y) = (\exists x)(\exists y)(x^2 = 2y^2),\ D_x = D_y = \mathbb{Z}$.

Solution.

(a)True

(b) False because $2 \leq 3 \longrightarrow \neg(2 \leq 3)$ is false .

(c) True because R(4,3) is true.

(d) False because $\sqrt{2}$ is irrational. ■

7. Negate the following statements

(a) All numbers are even

(b) For every integer x , either there exists integer y such that $x + y^2 = x^2$ or $x = 0$

8. Let $T(x)$, $C(x)$, and $S(x)$ mean "*x* is a triangle," "*x* is a circle," and "*x* is a square"; let $B(x)$, $G(x)$, and $Y(x)$ mean "*x* is blue," "*x* is gray," and "*x* is yellow"; let $RO(x, y)$, $AB(x, y)$, and $SC(x, y)$ mean "*x* is to the right of *y*," "*x* is above *y*," and "*x* has the same color as *y*"; and use the notation $x = y$ to denote the predicate "*x* is equal to *y*." Let the common domain *D* of all variables be the set of all the objects in the Tarski world. Use formal logical notation to write each of the following statements, and write a formal negation for each statement.

(a) For every circle *x*, *x* is above *f*.

(b) There is a square *x* such that *x* is yellow.

(c) For every circle *x*, there is a square *y* such that *x* and *y* have the same color.

(d) There is a square *x* such that for every triangle *y*, *x* is to the right of *y*.

Solution.

(a) *Statement*: $\forall x(C(x) \to AB(x,f))$

Negation:$\neg(\forall x\ (C(x) \to AB(x,f)))$

$\equiv \exists x \neg(C(x) \to AB(x,f))$

by the law for negating a ∀ statement

$\equiv \exists x(C(x) \land \neg AB(x,f))$

by the law of negating an if-then statement

(b) *Statement*: $\exists x(S(x) \land Y(x))$

Negation:$\neg(\exists x(S(x) \land Y(x)))$

$\equiv \forall x \neg(S(x) \land Y(x))$

by the law for negating a ∃ statement

$\equiv \forall x(\neg S(x)\ \lor \neg Y(x))$

by De Morgan's law

(c) *Statement*: $\forall x(C(x) \to \exists y(S(y) \land SC(x,y)))$

Negation: $\neg(\forall x(C(x) \to \exists y(S(y) \land SC(x,y))))$

$\equiv \exists x \neg(C(x) \to \exists y(S(y) \land SC(x,y)))$

by the law for negating a ∀ statement

$\equiv \exists\ x(C(x)\ \land\ \neg(\exists y(S(y) \land SC(x,y))))$

by the law for negating an if-then statement

$\equiv \exists x(C(x) \land \forall y(\neg(S(y) \land SC(x,y))))$

by the law for negating a ∃ statement

$$\equiv \exists x(C(x) \wedge \forall y(\neg S(y) \vee \neg SC(x,y)))$$

by De Morgan’s law

(d) *Statement:* $\exists x(S(x) \wedge \forall y(T(y) \rightarrow RO(x,y)))$

Negation: $\neg(\exists x(S(x) \wedge \forall y(T(y) \rightarrow RO(x,y))))$

$$\equiv \forall x \neg(S(x) \wedge \forall y(T(x) \rightarrow RO(x,y)))$$

by the law for negating a ∃ statement

$$\equiv \forall x(\neg S(x) \vee \neg(\forall y(T(y) \rightarrow RO(x,y))))$$

by De Morgan’s law

$$\equiv \forall x(\neg S(x) \vee \exists y \neg(T(y) \rightarrow RO(x,y))))$$

by the law for negating a ∀ statement

$$\equiv \forall x(\neg S(x) \vee \exists y(T(y) \wedge \neg RO(x,y))))$$

by the law for negating an if-then statement. ■

Exercise Set (2.2)

1- Let $P(x)$ denote the statement “$x \leq 4$”.

What are the truth values?

(a) $P(0)$; (b) $P(4)$; (c) $P(6)$.

2- Let $P(x)$ be the statement “*x* spends more than five hours every week day in class”, where the universe of discourse for *x* consists of all students.

Express each of these quantifications in ordinary language?

$$\exists x P(x);\ \forall x P(x);\ \exists x \neg P(x);\ \forall x \neg P(x).$$

3- Translate these statements into ordinary language, where $C(x)$ is "x is a comedian" and $F(x)$ is "x is funny" and the universe of discourse consists of all people.

(a)$\forall x(C(x) \rightarrow F(x))$;

(b)$\forall x(C(x) \wedge F(x))$;

(c)$\exists x(C(x) \rightarrow F(x))$;

(d)$\exists x(C(x) \wedge F(x))$.

4- Let $Q(x)$ be the statement "$x + 1 > 2x$". If the universe of discourse consists of all integers, what are these truths?

(a) $Q(0)$; (b) $Q(-1)$; (c) $Q(1)$; (d) $\exists x Q(x)$.

(e) $\forall x Q(x)$; (f) $\exists x \neg Q(x)$; (g) $\forall x \neg Q(x)$.

5- Suppose that the universe of discourse of the propositional function $P(x)$ consists of the integers 1, 2, 3, 4 and 5. Express these statements without using quantifiers, instead using only negation, disjunction and conjunction.

(a) $\exists x P(x)$; (b) $\forall x P(x)$; (c) $\neg \exists x P(x)$;

(d) $\neg\forall x P(x)$; (e) $\forall x((x \neq 3) \rightarrow P(x)) \vee \exists x \neg P(x)$.

6- Translate in two ways each of these statements into logical expressions using predicates, quantifiers and logical connectives. First, let the universe of discourse consist of the student in your class and second, let it consist of all people.

(a) Someone in your class can speak English;

(b) Everyone in your class is friendly;

(c) There is a person in your class who was not born in Assiut.

7- Translate the following statements into logical expressions using predicates, quantifiers and logical connectives.

(a) No one is perfect;

(b) Not everyone is perfect;

(c) All your friends are perfect;

(d) One of your friends is perfect.

8- Express each of the following statements using logical operators, predicate and quantifiers:

(a) Some propositions are tautologies.

(b) The negation of a contradiction is a tautology.

(c) The disjunction of two contingencies can be a tautology.

(d) The conjunction of two tautologies is a tautology.

9- Express each of these statements using quantifiers. Then form the negation of the statement, so that no negation is to before a quantifier. Next, express the negation in ordinary language (Do not simply use the words “It is not being case that”).

(a) Some old dogs can learn new tricks.

(b) No rabbit knows calculus.

(c) Every bird can fly.

(d) There is no dog that can talk.

(e) There is no one in this class who knows French and Russian.

10- Translate these statements into ordinary language, where the universe of discourse for each variable consists of all real numbers.

(a) $\forall x \exists y (x < y)$;

(b) $\forall x \forall y (x \geq 0 \wedge y \geq 0 \rightarrow xy \geq 0)$;

(c) $\forall x \forall y \forall z (xy = z)$.

11- Let $Q(x, y)$ be the statement "x has sent an e-mail message to y", where the universe of discourse for both x and y consists of all students in your class. Express each of these quantification in English.

(a) $\exists x \exists y Q(x, y)$;

(b) $\exists x \forall y Q(x, y)$.

12- Let $Q(x, y)$ be the statement "$x + y = x - y$". If the universe of discourse for both variables consists of all integers, what are the truth values of each of the following?

(a) $Q(1, 1)$; (b) $Q(2, 0)$; (c) $\forall y Q(1, y)$;

(d) $\exists x Q(x, 2)$; (e) $\exists x \exists y Q(x, y)$.

13- Determine the truth value of each of these statements if the universe of discourse of each variable consists of all real numbers.

(a) $\forall x \exists y (x^2 = y)$;

(b) $\forall x \exists y (x = y^2)$;

(c) $\exists x \forall y (xy = 0)$;

(d) $\exists x \exists y (x + y \neq y + x)$;

(e) $\forall x \big(x \neq 0 \rightarrow (x + y = 1)\big)$.

2.3 Rules of inference

◦ Arguments and Validity

Proofs in mathematics are valid arguments.

An argument is a sequence of statements that end with a conclusion.

♣ Many claims come to us as the conclusion of arguments. By "conclusion," we mean the claim that the argument is meant to defend.

♣ We will understand an argument as a finite list of logic forms (compound proposition), one of which is the conclusion, and the others are offered as reasons to believe the conclusion is true. We will call these other logic forms "premises." presented as follows:

$$\begin{array}{c} A_1 \\ A_2 \\ \vdots \\ \vdots \\ A_n \\ \hline \therefore\ B \end{array}$$

The logic forms above the bar are called premises while B is called the conclusion. (The symbol ∴ is read "therefore")

♣ We need to determine what makes an argument valid. Look at some examples.

(a) Premise 1: If Nixon was President, then Nixon was Commander in Chief.

Premise 2: Nixon was President.

Conclusion: Nixon was Commander in Chief.

(b) Premise 1: If Lincoln was an organism from deep in the sea, then Lincoln had three eyes.

Premise 2: Lincoln was an organism from deep in the sea.

Conclusion: Lincoln had three eyes.

(c) Premise 1: If Lincoln was President, then Lincoln had at least one portrait made of him.

Premise 2: Lincoln had at least one portrait made of him.

Conclusion: Lincoln was President.

♣ What is remarkable about the first argument is that if the premises are true it seems the conclusion must be true. That is an excellent standard to have for arguments, since it describes a clear relation between premises and a conclusion.

A valid argument

Definition.

A valid argument is an argument in which if the premises are true then the conclusion must be true.

♣ Note that the second argument is absurd. Both premises and the conclusion are false. Does that mean it is a bad argument.?

Well, we could define a bad argument to be one where all the logic forms are false, but this would confuse the structure of the argument with the truth value of the logic forms that compose it. Our interest, right now, is argumentation itself. In that case, we must recognize that the second argument is **valid**. If the premises were true, the conclusion would have to be true. Valid arguments can have false conclusions if some of their premises are false.

♣ It is useful, therefore, to distinguish valid arguments with true premises from valid arguments with some false premises. We will call arguments like the first argument above "**sound**."

A sound argument

Definition.

A sound argument is an argument which is valid and which has true premises. An argument that is not sound is called **unsound**.

♣ Note that the third argument, even though each logic form in it is true, is **invalid**. It is not the case that if one has a portrait one was President. But that is the reasoning that underlies the leap from premise 2 to the conclusion. Be aware that invalid arguments can contain all true statements. They are invalid because other arguments of the exact same form could have true premises and a false conclusion.

● Propositional Logic and Connectives

♣ Logic is a formal method which provides a way to rigorously test arguments for validity. We will look at one part of logic -- propositional logic -- in order to illustrate and clarify the nature of validity and good reasoning in arguments.

♣ Propositional logic is formulated out of propositions and "connectives." Connectives are ways of putting propositions together to make new propositions.

♣ We will represent propositions with letters *p*, *q*, *r*, ...

♣ Thus, the 3 arguments above could be represented:

(a) Premise 1: If *p*, then *q*.

Premise 2: *p*.

Conclusion: *q*.

(b) Premise 1: If *r*, then *s*.

Premise 2: *r*.

Conclusion: *s*.

(c) Premise 1: If *t*, then *v*.

Premise 2: *v*.

Conclusion: *t*.

Assuming that we interpret our letters to be standing for the propositions:

- *p*: Nixon was President
- *q*: Nixon was Commander in Chief.
- *r*: Lincoln was an organism from deep in the sea.
- *s*: Lincoln had three eyes.
- *t*: Lincoln was President.
- *v*: Lincoln had at least one portrait made of him.

These arguments use the connective "if ... then ....".

To further abbreviate our logic, we will replace these English words with a single arrow: →.

Thus, instead of "If p, then q" we will write "$p \rightarrow q$".

Our first argument would then look like:

Premise 1: $p \rightarrow q$.

Premise 2: p.

Conclusion: q.

●Methods to Test Validity of an Argument

First Method to prove validity

Definition.

A valid argument is a finite set of propositions $p_1,\ p_2, \ldots, p_r$ (premises), together with a proposition c, the conclusion, such that the propositional form $(p_1 \wedge\ p_2 \wedge \ldots \wedge p_r) \rightarrow c$ is a tautology.

We say c *follows logically from*, or is *a logical consequence* of the premises.

We write $p_1,\ p_2, \ldots, p_r \vdash c$. The symbol ⊢ is called the *turnstile*.

Example.

Determine whether the following argument is valid or invalid.

Premises 1: "if you have a current password, then you can log onto the network".

Premises 2: "you have a current password".

Therefore

Conclusion: "you can log onto the network".

Solution.

Let p represent: "you have a current password"

and

q represent: "you can log onto the network"

Then the argument has the form

$$\begin{array}{c} p \longrightarrow q \\ \underline{\quad p \quad} \\ \therefore\; q \end{array}$$

When p and q are proposition variables, the statement $[(p \to q) \land p] \to q$ is a tautology.

In this case we say that this argument is valid.■

Example.

Let $p_1 =$ "John graduates"

$p_2 =$ "Mary graduates"

$p_3 =$ "John gets a job"

$p_4 =$ "Mary gets a job"

$p_5 =$ "Mary earns money"

(i) Consider the following argument:

"If John graduates then he gets a job".

"John graduates".

"Therefore John gets a job".

To see the "form" of this argument we symbolize it as

$$p_1 \to p_3, p_1 \vdash p_3$$

Now, the student can prove that $\big((p_1 \to p_3) \land p_1\big) \to p_3$ is a tautology. So, the given argument is valid.

Note Another "instance" of this argument follows if we set $p_1 =$ "$2 < 1$" and $p_3 =$ "$3 < 2$".

The argument then reads:

"If $2 < 1$ then $3 < 2$".

"$2 < 1$".

"Therefore, 3 < 2".

This is still valid though some of the propositions, i.e. $2 < 1$ for instance, are false.

(ii) Consider the following argument:

“If Mary graduates then she gets a job”.

“Mary does not get a job”.

“Therefore Mary does not graduate”.

Symbolized, this becomes

$$p_2 \to p_4, (\neg\, p_4) \vdash (\neg p_2).$$

Since $\big((p_1 \to p_4) \wedge (\neg p_4)\big) \to (\neg p_2)$ is a tautology, then the given argument is valid.

(iii) Consider the following argument:

“Either Mary or John graduate”.

“John does not graduate”.

“Therefore Mary graduates”.

Symbolized, this becomes $p_2 \vee p_1, (\neg\, p_1) \vdash p_2$.

Since $[(p_2 \vee p_1) \wedge (\neg\, p_1)] \to p_2$ is a tautology, then the given argument is valid.

(iv) Consider the following argument:

“If Mary graduates then she gets a job”.

“If Mary gets a job then she earns money”.

“Therefore if Mary graduates then she earns money”.

Symbolized, this becomes

$$p_2 \to p_4 , p_4 \to p_5 \vdash \; p_2 \to p_5.$$

Now, $[(p_2 \to p_4) \; \wedge (p_4 \to p_5)] \to \; (p_2 \to p_5)$ is a tautology. Then the given argument is valid.■

We can sum up the above by saying the following are all valid:

(i) $p \to q, p \vdash q$;

(ii) $p \to q, (\neg\, q) \vdash (\neg\, p)$;

(iii) $p \vee q, (\neg\, q) \vdash p$;

(iv) $p \to q \;, q \to r \vdash \; p \to r$.

Example.

Show that $p \to q, p \vee q \vdash (\neg\, p) \vee (\neg\, q)$ is invalid.

p	q	$p \to q$	$p \vee q$	$(p \to q) \wedge (p \vee q)$ $\boldsymbol{P}$	$\neg\, p \vee \neg\, q$ $\boldsymbol{Q}$	$\boldsymbol{P \to Q}$
1	1	1	1	1	0	**0**
1	0	0	1	0	1	**1**
0	1	1	1	1	1	**1**
0	0	1	0	0	1	**1**

We do not have a tautology in the last column so the argument is **invalid**. ■

Second Method to prove validity

Note If an argument is valid then $(p_1 \wedge\ p_2 \wedge \ldots \wedge p_r) \rightarrow c$ is a tautology, and so it is always true. So we need to prove that it is **never** false. It can only be only false if c is false and $p_1 \wedge\ p_2 \wedge \ldots \wedge p_r$ is true, i.e. all $p_1,\ p_2, \ldots, p_r$ are true. So, we **never** want to see a row in the truth table where all the premises are true and the conclusion false. This observation gives a second way of checking that an argument is valid or not.

To check that an argument is valid or not we do the following steps.

1. Identify the premises and conclusion of the argument.

2. Construct a truth table showing the truth values of all the premises and conclusion.

3. Find the rows (**called critical rows**) in which all the premises are true.

4. In each critical row, determine whether the conclusion of the argument is also true.

(a) If in each critical row the conclusion is also true, then the argument form is **valid**.

(b) If there is at least one critical row in which conclusion is false, the argument form is **fallacy** (**invalid**).

Example.

Determine whether the following argument is valid or invalid.

$$\begin{array}{c} p \to q \\ \underline{\quad p \quad} \\ \therefore q \end{array}$$

Solution.

The truth table for the premises and conclusion is:

p	q	$p \to q$
1	1	1
1	0	0
0	1	1
0	0	1

The **first** line is the only **critical line**, where the premises is true. We see that the conclusion is also true. Then the given argument is valid. ■

Example.

Is $p \to q, p \vee q \vdash (\neg p) \vee (\neg q)$ is valid?

Solution.

We look at the truth table:

p	q	$p \to q$	$p \vee q$	$(p \to q) \wedge (p \vee q)$ $\boldsymbol{P}$	$\neg p \vee \neg q$ $\boldsymbol{Q}$	$\boldsymbol{P \to Q}$
1	1	1	1	1	0	**0**
1	0	0	1	0	1	**1**
0	1	1	1	1	1	**1**
0	0	1	0	0	1	**1**

In the first line the conclusion is false, but all premises are true. Hence the argument is **invalid**.

This method requires fewer columns than in the first method. ■

Is there an "instance" of the above argument which is "obviously" invalid?

Try looking in the "World of Mathematics", for instance, choosing $p \equiv$ "$3 > 2$" and $q \equiv$ "$2 > 1$".

Then the argument becomes:

If $3 > 2$ then $2 > 1$,

Either $3 > 2$ or $2 > 1$,

Therefore, either $3 \leq 2$ or $2 \leq 1$.

Both premises are true but the conclusion is false. On the basis that we never want a false conclusion to follow from true premises, this argument is invalid.

But be careful! Consider another instance.

So let p $\equiv$ "Assiut is a city" and q $\equiv$ "Suhag is a city".

Then the argument becomes:

"If Assiut is a city then Suhag is a city".

"Either Suhag or Assiut is a city".

"Therefore, either Suhag is not a city or Assiut is not a city".

If I tell you that Suhag is a city but Assiut is not a city then you can check that all the propositions in this argument are true. But the argument is still invalid. It is a case of the conclusion, though true, not following logically from the true premises.

Example.

Is $p \rightarrow (s \rightarrow (\neg r)), p \rightarrow r, p \vdash \neg s$ is valid?

Solution.

We look at the truth table:

p	r	s	$\neg r$	$s \rightarrow \neg r$	$p \rightarrow (s \rightarrow \neg r)$	$p \rightarrow r$	p	$\neg s$
1	1	1	0	0	0	1	1	0
1	1	0	0	1	**1**	**1**	**1**	**1**
1	0	1	1	1	1	0	1	0
1	0	0	1	1	1	0	1	1
0	1	1	0	0	1	1	0	0
0	1	0	0	1	1	1	0	1
0	0	1	1	1	1	1	0	0
0	0	0	1	1	1	1	0	1

We look at each row in turn. We look to see if on any row we have a case of all the premises being true with the conclusion false. For instance in the first row we see that premises are 0, 1, 1 and the conclusion 0. This is allowable. By checking each row we see that each row is allowable, that is, we never have a case of all premises true with the conclusion false.

Hence the argument is **valid**. ■

Example.

Determine if the next arguments is valid or invalid.

(1) $p \to q, p \to r \vdash p \to (q \wedge r)$;

(2) $p \to q, r \to s, \neg q \vee \neg s, r \vee \neg q \vdash p \leftrightarrow \neg r$.

Solution.

We use the critical lines where in (2) we need to construct $2^4 \ = 16$ lines.

(1) We look at the truth table:

p	q	r	$q \wedge r$	$p \to q$	$p \to r$	$p \to (q \wedge r)$
1	**1**	**1**	**1**	**1**	**1**	**1**
1	1	0	0	1	0	0
1	0	1	0	0	1	0
1	0	0	0	0	0	0
0	**1**	**1**	**1**	**1**	**1**	**1**
0	**1**	**0**	**0**	**1**	**1**	**1**
0	**0**	**1**	**0**	**1**	**1**	**1**
0	**0**	**0**	**0**	**1**	**1**	**1**

Here the critical lines are (1), (5), (6), (7), and (8). The conclusion is true in all of these lines. So, the argument is valid.

(2) (Exercise for the student).

Note I have given here two methods for using a truth table to check whether $p_1,\ p_2, \dots, p_r \vdash c$ is valid or not. Do not *mix* up these methods!

In the *first method* use a truth table to work out the truth values of $(p_1 \wedge\ p_2 \wedge \dots \wedge p_r) \rightarrow c$, and hope that it is always true, i.e. a tautology.

In the *second method* construct a table containing a column for each of the $p_1,\ p_2, \dots$ up to p_r along with c and hope that there is no row with all the p_i true and c false.

The second method of proving validity needs a smaller number of columns than the first, but if the number of basic propositions p, q, r, etc. is large then the tables in both methods need a large number of rows. Thus the tables get cumbersome in both methods and an alternative method is necessary (Rules of inference of propositional logic and quantifiers).

If an argument is **invalid** there is sometimes a quick method of showing this.

Example.

Show that $(p \vee q) \rightarrow s, q \rightarrow s \vdash s$ is invalid.

Solution.

We do this by trying to make the conclusion false and the premises all true.

The conclusion if false if we choose s to be false. Then $q \rightarrow s$ can be true only if q is false.

Finally, for $(p \vee q) \rightarrow s$ to be true we require $p \vee q$ to be false, and so p must be false.

Hence if all of p, q and s are false (i.e. the bottom row of the truth table), we see that all the premises are true but the conclusion is false. Hence the argument is invalid. ■

Note.

To determine whether the argument which contains n variables is valid or invalid we need 2^n lines. It is difficult to use truth table for large n. So, we use the definition of the valid argument. We find a critical line with false conclusion. If not we have a valid argument.

Example.

Determine if the next arguments is valid or invalid.

(a) $p \rightarrow q, q \rightarrow (p \rightarrow r), p \vdash p \rightarrow (q \wedge r)$;

(b) $p \rightarrow q, r \rightarrow \neg p, r \rightarrow q \vdash q$.

Solution.

(a) Exercise.

(b) Starting with the conclusion and assume q is 0. The first premise $p \rightarrow q$ is 1 when p is 0. So, the third premise $r \rightarrow q$ is 1 when r is 0. This implies that the second premise $r \rightarrow \neg p$ is 1. Therefore the given argument is **invalid**.

We have obtained the following critical line:

p	q	r	$p \rightarrow q$	$r \rightarrow \neg p$	$r \rightarrow q$	q
0	0	0	1	1	1	0

■

● Rules of Inference for Propositional Logic

We can always use a truth table to show that an argument form is valid. We do this by showing that whenever the premises are true, the conclusion must also be true. However, this can be a tedious approach. For example, when an argument form involves 10 different propositional variables, to use a truth table to show this argument form is valid requires $2^{10} = 1024$ different rows. Fortunately, we do not have to resort to truth tables. Instead, we can first establish the validity of some relatively simple argument forms, called rules of inference. These rules of inference can be used as building blocks to construct more complicated valid argument forms. We will now introduce the most important rules of inference in propositional logic.

The tautology $(p \wedge (p \rightarrow q)) \rightarrow q$ is the basis of the rule of inference called **modus ponens** (Law of Detachment). This Latin term means "**Method of affirming**" (since the conclusion is an affirmation). This tautology leads to the following valid argument form, which we have already seen in our initial discussion about arguments:

$$\begin{array}{c} p \\ p \to q \\ \hline \therefore q \end{array}$$

In particular, modus ponens tells us that if a conditional statement and the hypothesis of this conditional statement are both true, then the conclusion must also be true.

Example.

Suppose that the conditional statement “"if $n > 3$, then $n^2 > 9$" is true , consequently if $n > 3$, then by modus ponens $n^2 > 9$.■

Example.

Suppose that the conditional statement "If it snows today, then we will go skiing" and its hypothesis, "It is snowing today," are true. Then, by modus ponens, it follows that the conclusion of the conditional statement, "We will go skiing" is true. ■

As we mentioned, a valid argument can lead to an incorrect conclusion if one or more of its premises is false. We illustrate this again in the following example.

Example.

Determine whether the argument given here is valid and determine whether its conclusion must be true because of the validity of the argument.

"If $\sqrt{2} > \frac{3}{2}$, then$\left(\sqrt{2}\right)^2 > \left(\frac{3}{2}\right)^2$. We know that $\sqrt{2} > \frac{3}{2}$.

Consequently, $\left(\sqrt{2}\right)^2 = 2 > \left(\frac{3}{2}\right)^2 = \frac{9}{4}$."

Solution.

Let p be the proposition "$\sqrt{2} > \frac{3}{2}$" and q the proposition "$2 > \left(\frac{3}{2}\right)^2$". The premises of the argument are $p \rightarrow q$ and p, and q is the conclusion. This argument is valid because it is constructed by using modus ponens, a valid argument form. However, one of its premises $\sqrt{2} > \frac{3}{2}$ is false. Consequently, we cannot conclude that the conclusion is true. Furthermore, note that the conclusion of this argument is false, because $2 < \frac{9}{4}$. ■

The following table lists some important rules of inference

Rule of inference	Tautology	Name
$\begin{array}{c} p \\ \hline \therefore p \vee q \end{array}$	$p \to p \vee q$	Addition
$\begin{array}{c} p \wedge q \\ \hline \therefore p \end{array}$	$p \wedge q \to p$	Simplification
$\begin{array}{c} p \\ q \\ \hline \therefore p \wedge q \end{array}$	$p \wedge q \to p \wedge q$	Conjunction
$\begin{array}{c} p \\ p \to q \\ \hline \therefore q \end{array}$	$(p \wedge (p \to q)) \to q$	Modus ponens
$\begin{array}{c} \neg q \\ p \to q \\ \hline \therefore \neg p \end{array}$	$(\neg q \wedge (p \to q)) \to \neg p$	Modus tollens
$\begin{array}{c} p \to q \\ q \to r \\ \hline \therefore p \to r \end{array}$	$\big((p \to q) \wedge (q \to r)\big) \to (p \to r)$	Hypothetical syllogism
$\begin{array}{c} p \vee q \\ \neg p \\ \hline \therefore q \end{array}$	$\big((p \vee q) \wedge \neg p\big) \to q$	Disjunction syllogism
$\begin{array}{c} p \vee q \\ \neg p \vee r \\ \hline \therefore q \vee r \end{array}$	$\big((p \vee q) \wedge (\neg p \vee r)\big) \to (q \vee r)$	Resolution

Exercise

Which rule of inference is used in each argument below?

- ☺“Alice is a Math major”.
  Therefore, “Alice is either a Math major or a CSI major”.
- ☺“Jerry is a Math major and a CSI major”.
  Therefore, “Jerry is a Math major”.
- ☺“If it is rainy, then the pool will be closed”. “It is rainy”.
  Therefore, “the pool is closed”.
- ☺“If it snows today, the university will close”. “The university is not closed today”.
  Therefore, “it did not snow today”.
- ☺“If I go swimming, then I will stay in the sun too long”.
  “If I stay in the sun too long, then I will sunburn”.
  Therefore, “if I go swimming, then I will sunburn”.
- ☺“I go swimming or eat an ice cream”. “I did not go swimming”.”
  Therefore, “I eat an ice cream”.

Example.

State which rule of inference is the basis of the following argument

"It is below freezing now".

Therefore, "it is either below freezing or raining now".

Solution.

Let p be the proposition "It is below freezing now" and let q be the proposition "It is raining now". Then this argument is of the form

$$\frac{p}{\therefore p \vee q}$$

This argument uses the **addition** rule. ■

Example.

State which rule of inference is used in the argument:

"If it rains today, then we will not have a barbecue today.

"If we do not have a barbecue today, then we will have a barbecue tomorrow."

Therefore, "if it rains today, then we will have a barbecue tomorrow."

Solution.

Let *p* be the proposition “It is raining today”, let *q* be the proposition “we will not have a barbecue today” and let *r* be the proposition “we will have a barbecue tomorrow”. Then this argument is of the form:

$$\begin{array}{c} p \to q \\ q \to r \\ \hline \therefore p \to r \end{array}$$

Hence, this argument is **a hypothetical syllogism**. ■

Example.

State which rule of inference is the basis of the following argument:

“It is below freezing and raining now”.

Therefore, “It is below freezing now.”

Solution.

Let *p* be the proposition “It is below freezing now,” and let *q* be the proposition “It is raining now.” This argument is of the form

$$\begin{array}{c} p \wedge q \\ \hline \therefore p \end{array}$$

This argument uses the **simplification** rule. ■

Example.

State which rule of inference is the basis of the following argument:

"If Zeus is human, then Zeus is mortal."

"Zeus is not mortal."

Therefore, "Zeus is not human."

Solution.

Let *p* be the proposition "Zeus is human," and let *q* be the proposition "Zeus is mortal." This argument is of the form

$$\begin{array}{c} p \to q \\ \neg q \\ \hline \therefore \neg p \end{array}$$

The fact that this argument is valid is called **Modus Tollens** which means (**Method of denying**) since the conclusion is denial. ■

● Using Rules of Inference to Build Arguments

When there are many premises, several rules of inference are often needed to show that an argument is valid.

Example.

Show that the hypotheses:

▶It is not sunny this afternoon and it is colder than yesterday.

▶We will go swimming only if it is sunny.

▶If we do not go swimming, then we will take a canoe trip.

▶If we take a canoe trip, then we will be home by sunset.

Lead to the conclusion:

▶We will be home by sunset.

Solution.

Main steps:

1. Translate the statements into proposional logic.

2. Write a formal proof, a sequence of steps that state hypotheses or apply inference rules to previous steps.

Assume the following propositions:

p**:** it is sunny this afternoon

q**:** it is colder than yesterday

r: we will go swimming

s: we will take a canoe trip

t: we will be home by sunset

Then the hypotheses are $\neg p \wedge q, r \longrightarrow p$, $\neg r \longrightarrow s$, $s \longrightarrow t$, and the conclusion is simply t.

We construct an argument to show that desired conclusion as follows:

Step	Reason
(1) $\neg p \wedge q$	Hypothesis
(2) $\neg p$	Simplification using step (l).
(3) $r \longrightarrow p$	Hypothesis
(4) $\neg r$	Modus tollens using (2) and (3)
(5) $\neg r \rightarrow s$	Hypothesis
(6) s	Modus ponens using (4) and (5)
(7) $s \longrightarrow t$	Hypothesis
(8) t Conclusion	Modus ponens using (6) and (7)

Example.

Using rules of valid inference to solve the problem:

(a) If my glasses are on the kitchen table, then I saw them at breakfast.

(b) I was reading the newspaper in the living room or I was reading in the kitchen.

(c) If I was reading the newspaper in the living room. then my glasses are on the coffee table.

(d) I did not see my glasses at breakfast.

(e) If I was reading my book in bed, then my glasses are on the bed table.

(f) If I was reading the newspaper in the kitchen, then my glasses are on the kitchen table.

Where are the glasses?

Solution.

p : my glasses are on the kitchen table.

q : I saw them at breakfast.

r : I was reading the newspaper in the living room.

s : I was reading the newspaper in the kitchen.

t : my glasses are on the coffee table.

u : I was reading my book in bed.

v : my glasses are on the bed table.

Then the given statements are:

(a) $p \rightarrow q$; (b) $r \vee s$; (c) $r \rightarrow t$;

(d) $\neg q$; (e) $u \rightarrow v$; (f) $s \rightarrow p$.

We construct an argument to show that desired conclusion as follows:

Step	Reason
(l) $p \rightarrow q$	Hypothesis (a)
(2) $\neg q$	Hypothesis (d)
(3) $\therefore \neg p$	Modus tollens using (1) and (2)
(4) $s \rightarrow p$	Hypothesis (f)
(5) $\neg p$	Conclusion (3)
(6) $\therefore \neg$s	Modus tollens using (4) and (5)
(7) $r \vee s$	Hypothesis (b)
(8) $\therefore r$	disjunctive syllogism using (6) and (7)
(9) $r \rightarrow t$	Hypothesis (c)
(10) t	Modus Ponens using (8) and (10)

Hence t is true and the glasses are on the coffee table.■

●Resolution and Automated Theorem Proving

We can build programs that automate the task of reasoning and proving theorems.

Many of these programs make use of a rule of inference known as **a resolution**. This rule of inference is based on the tautology $\big((p \vee q) \wedge (\neg p \vee r)\big) \rightarrow (q \vee r)$.

The final disjunction in the resolution rule, $q \vee r$, is called the **resolvent**.

If we express the hypotheses and the conclusion as clauses (possible by CNF, a conjunction of clauses), we can use resolution as the only inference rule to build proofs!

This is used in programming languages like Prolog.

It can be used in automated theorem proving systems.

Example.

Use resolution to show that the hypothesis:

"Ahmed is skiing or it is not snowing"

and

"It is snowing or Ali is playing hockey"

imply that "Ahmed is skiing or Ali is playing hockey".

Solution.

p: it is snowing

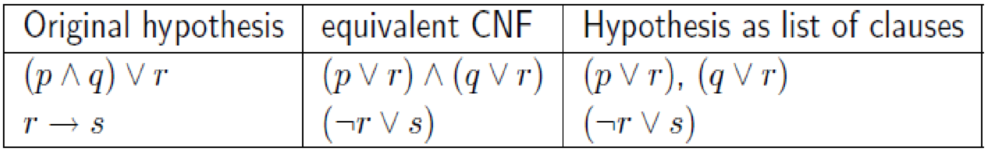
q: Ahmed is skiing

r: Ali is playing hockey

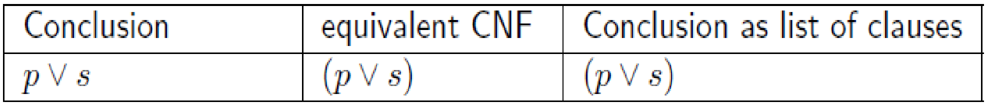
We can represent the hypothesis as $\neg p \vee q$ and $p \vee r$, respectively. Using resolution, the proposition $q \vee r$ "Ahmed is skiing or Ali is playing hockey" follows.■

⁎Proofs that use exclusively resolution as inference rule

Step 1: Convert hypotheses and conclusion into clauses:

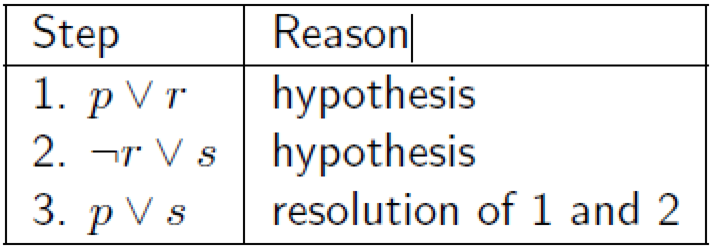


Original hypothesis	equivalent CNF	Hypothesis as list of clauses
$(p \wedge q) \vee r$	$(p \vee r) \wedge (q \vee r)$	$(p \vee r)$, $(q \vee r)$
$r \rightarrow s$	$(\neg r \vee s)$	$(\neg r \vee s)$



Conclusion	equivalent CNF	Conclusion as list of clauses
$p \vee s$	$(p \vee s)$	$(p \vee s)$

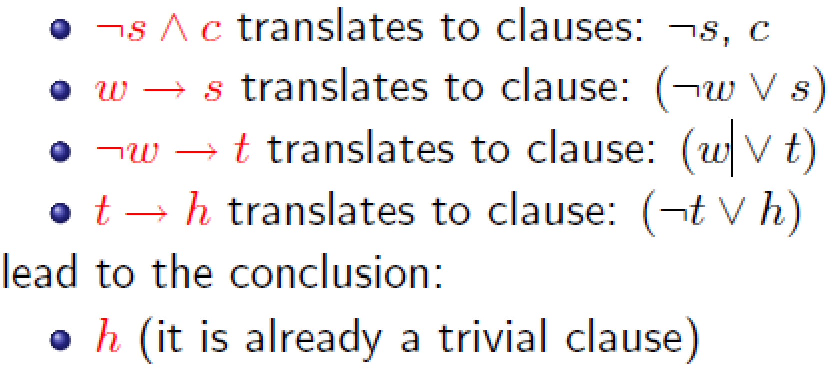
Step 2: Write a proof based on resolution:



Step	Reason
1. $p \vee r$	hypothesis
2. $\neg r \vee s$	hypothesis
3. $p \vee s$	resolution of 1 and 2

Example.

Show that the hypotheses:



- $\neg s \wedge c$ translates to clauses: $\neg s$, c
- $w \longrightarrow s$ translates to clause: $(\neg w \vee s)$
- $\neg w \longrightarrow t$ translates to clause: $(w \vee t)$
- $t \longrightarrow h$ translates to clause: $(\neg t \vee h)$

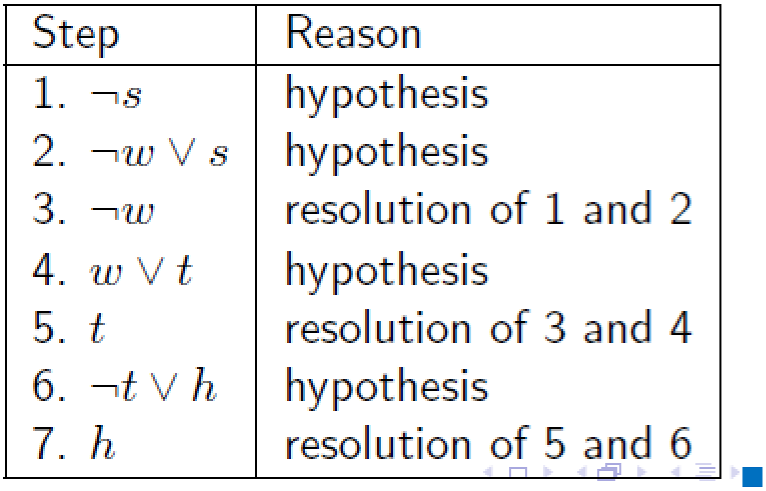
lead to the conclusion:

- h (it is already a trivial clause)

Note that the fact that p and $\neg p \vee q$ implies q (disjunctive syllogism) is a special case of resolution, since $p \vee 0$ and $\neg p \vee q$ give us $0 \vee q$ which is equivalent to q.

Proof.

Resolution-based proof:



Step	Reason
1. $\neg s$	hypothesis
2. $\neg w \vee s$	hypothesis
3. $\neg w$	resolution of 1 and 2
4. $w \vee t$	hypothesis
5. t	resolution of 3 and 4
6. $\neg t \vee h$	hypothesis
7. h	resolution of 5 and 6

●**Fallacies**

Fallacy = misconception resulting from incorrect argument.

►Fallacy of affirming the conclusion

Based on

$$((p \to q) \land q) \to p$$

which is NOT A TAUTOLOGY.

Example.

If prof gives chocolate, then you answer the question. You answer the question. We conclude the prof gave chocolate.■

►Fallacy of denying the hypothesis

Based on

$$((p \to q) \land \neg p) \to \neg q$$

which is NOT A TAUTOLOGY.

Example.

If prof gives chocolate, then you answer the question. Prof doesn't give chocolate. Therefore, you don't answer the question. ■

● Rules of Inferences for Quantified Statements

We have discussed rules of inference for propositions. Now, we will describe some important rules of inference for statements involving quantifiers. These rules of inference are used extensively in mathematics arguments, often without being explicitly mentioned.

♣ **Universal instantiation** is the rule of inference used to conclude that $P(c)$ is true, where c is a particular member of the domain, given the premise $\forall x P(x)$. Universal instantiation is used when we conclude from the statement "All women are wise" that "Huda is wise", where Hodi is a member of the universe of discourse of all women.

♣ **Universal generalization** is the rule of inference that states that $\forall x P(x)$ is true, given the premise that $P(c)$ is true for all element c in the domain. Universal generalization is used when we show that $\forall x P(x)$ is true by taking an arbitrary element c from the domain and showing that $P(c)$ is true. The element c that we select must be arbitrary, and not a specific, element of the domain. That is, when we assert from $\forall x P(x)$ the existence of an element c in the domain, we have no control over c and cannot make any other assumptions about c other than it comes from the domain. Universal generalization is used implicitly in many proofs in mathematics and is seldom mentioned explicitly. However, the error of adding unwarranted assumptions about the arbitrary element c when universal generalization is used is all too common in incorrect reasoning.

♣**Existential instantiation** is the rule that allows us to conclude that there is an element c in the domain for which $P(c)$ is true if we know that $\exists x P(x)$ is true.

We cannot select an arbitrary value of c here, but rather it must be a c for which $P(c)$ is true. Usually we have no knowledge of what c is, only that it exists. Because it exists, we may give it a name (c) and continue our argument.

♣**Existential generalization** is the rule of inference that is used to conclude that $\exists x P(x)$ is true when a particular element c with $P(c)$ true is known. That is, if we know one element c in the domain for which $P(c)$ is true, then we know that $\exists x P(x)$ is true.

We summarize these rules of inference for statement.

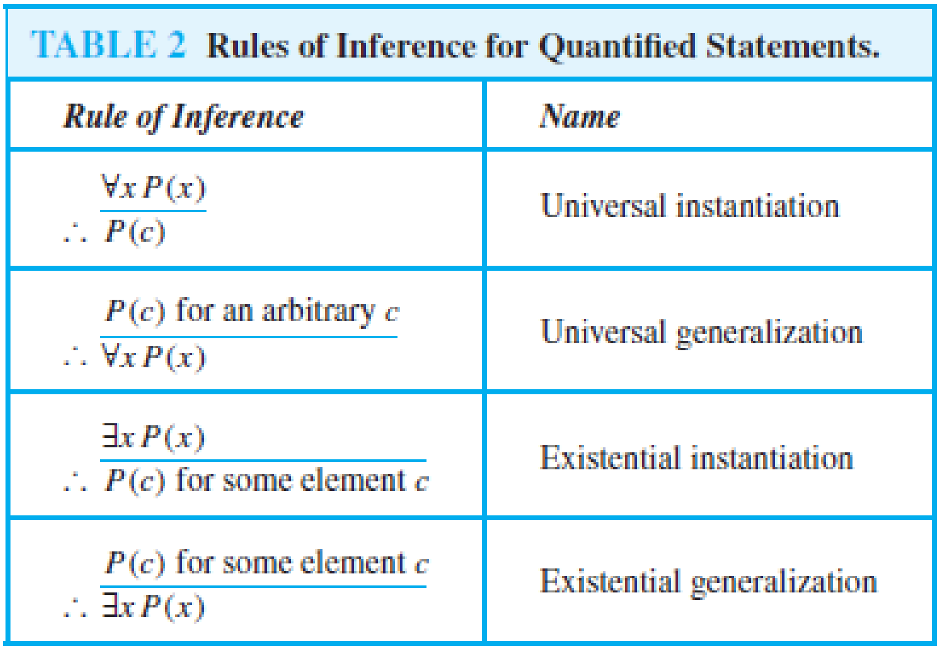


TABLE 2 Rules of Inference for Quantified Statements.

Rule of Inference	*Name*
$\forall x P(x)$ $\therefore P(c)$	Universal instantiation
$P(c)$ for an arbitrary c $\therefore \forall x P(x)$	Universal generalization
$\exists x P(x)$ $\therefore P(c)$ for some element c	Existential instantiation
$P(c)$ for some element c $\therefore \exists x P(x)$	Existential generalization

We will illustrate how some of these rules of inference for quantified statements are used in the following examples.

Example.

State which rule of inference is applied in the following argument.

Let *c* be any student.

“Student *c* has a personal computer”.

Therefore, “all student has a personal computer”.

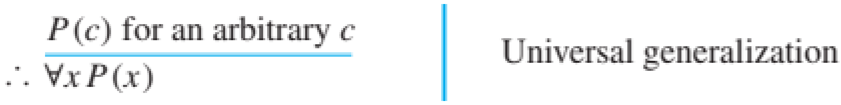
Solution.

Determine individual propositional function

$P(x)$: x has a personal computer.

Domain: all students.

The argument using $P(x)$:



$$\begin{array}{l|l} \underline{\quad P(c) \text{ for an arbitrary } c\quad} & \text{Universal generalization} \\ \therefore \forall x P(x) & \end{array}$$

Domain: all students

(c is an arbitrary element of the domain.)

Example.

Show if $\forall x\ (P(x) \wedge Q(x))$ is true then $\forall x P(x) \wedge \forall x Q(x)$ is true. (using direct technique)

Solution.

Assume $\forall x\ (P(x) \wedge Q(x))$ is true.

If a is in the domain then $P(a) \wedge Q(a)$ is true by universal instantiation.

So, $P(a)$ is true and $Q(a)$ is true.

Element a can be any element in the domain.

So, $\forall x P(x)$ is true and $\forall x P(x)$ is true by universal generalization.

Thus, $\forall x P(x) \wedge \forall x Q(x)$ is true. ■

Example.

State which rule of inference is applied in the argument.

There is a person in the store.

Therefore, some person *c* is in the store.

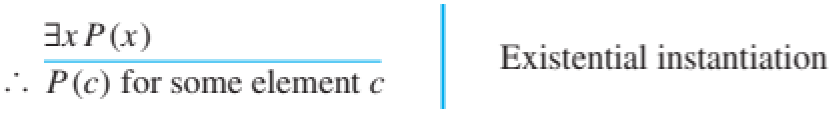
Solution.

Determine individual propositional function

$P(x)$: x is in the store.

Domain: all people

The argument using $P(x)$:



$\exists x P(x)$ $\therefore P(c)$ for some element c	Existential instantiation

(c is some element of the domain.)

Example.

State which rule of inference is applied in the argument.

His dog is playing in the park.

Therefore, there is a dog playing in the park.

Solution.

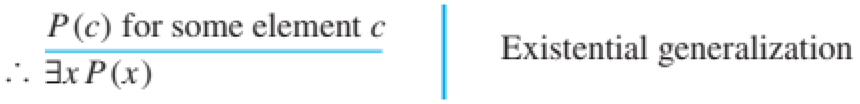
Determine individual propositional function

$P(x)$: x is playing in the park.

c: his dog

Domain: all dogs

The argument using $P(x)$.



$P(c)$ for some element c $\therefore \exists x P(x)$	Existential generalization

.■

Example.

Show if $\exists x\ (P(x) \wedge Q(x))$ is true then $\exists x P(x) \wedge \exists x Q(x)$ is true. (using direct technique)

Solution.

Assume $\exists x\ (P(x) \wedge Q(x))$ is true.

Let a be some element of the domain, that $P(a) \wedge Q(a)$ is true by existential instantiation.

So, $P(a)$ is true and $Q(a)$ is true.

So, $\exists x P(x)$ is true and $\exists x P(x)$ is true by Existential generalization.

Thus, $\exists x P(x) \wedge \exists x Q(x)$ is true. ■

Example.

Show that the premises "Everyone in this discrete mathematics class has taken a course in computer science" and "Aly is a student in this class" imply the conclusion "Aly has taken a course in computer science".

Solution.

$D(x)$: x is in the discrete math class"

$C(x)$: x has taken a course in computer science"

The argument:

$$\begin{array}{c} \forall x(D(x) \rightarrow C(x)) \\ D(\text{Aly}) \\ \hline \therefore\ C(\text{Aly}) \end{array}$$

Then the premises are $\forall x\big(D(x) \rightarrow C(x)\big)$ and $D(\text{Aly})$

The conclusion is $C(\text{Aly})$

The following steps can be used to establish the conclusion from the premises.

Step	Reason
1. $\forall x\big(D(x) \rightarrow C(x)\big)$	Premise
2. $D(\text{Aly}) \rightarrow C(\text{Aly})$	Universal instantiation by 1
3. $D(\text{Aly})$	Premise
4. $C(\text{Aly})$	Modus ponens from 2 and 3

.■

Example.

Show that the premises:

“A student in this class has not read the book” and “Everyone in this class passed the first exam.”

imply the conclusion:

“Someone who passed the first exam has not read the book”.

Solution.

$C(x)$: x is in this class

$B(x)$: x has read the book

$P(x)$: x passed the first exam.

The premises:

$$\exists x\big(C(x) \wedge \neg B(x)\big).$$

And

$$\forall x\big(C(x) \rightarrow p(x)\big).$$

The conclusion:

$$\exists x\big(p(x) \wedge \neg B(x)\big).$$

The following steps can be used to establish the conclusion from the premises.

Step	Reason
(1) $\exists x\big(C(x) \land \neg B(x)\big)$	Premise
(2) $C(a) \land \neg B(a)$	Existential instantiation from (1)
(3) $C(a)$	Simplification form (2)
(4) $\forall x\big(C(x) \rightarrow P(x)\big)$	Premise
(5) $C(a) \rightarrow P(a)$	Universal instantiation from (4)
(6) $P(a)$	Modus ponens from (3)and (5)
(7) $\neg B(a)$	Simplification of (2)
(8) $P(a) \land \neg B(a)$	Conjunction from (6) and (7)
(9) $\exists x\big(P(x) \land \neg B(x)\big)$	Existential generalization from (8)

. ■

●Combining Rules of Inference for Propositions and Quantified Statements

These inference rules are frequently used and combine propositions and quantified statements:

●Universal Modus Ponens

$$\begin{array}{rl} & \forall x(P(x) \rightarrow Q(x)) \\ & P(a) \text{, where } a \text{ is a particular element in the domain} \\ \hline \therefore & Q(a) \end{array}$$

Universal modus ponens is commonly used in mathematical arguments. This is illustrated by the following example.

Example.

Assume that “For all positive integers *n*, if *n* is greater than 4, then n^2 is less than 2^n” is true.

Use universal modus ponens to show that $100^2 < 2^{100}$.

Solution.

Let $P(n)$ denote “$n > 4$” and $Q(n)$ denote “$n^2 < 2^n$.”

The statement “For all positive integers *n*, if *n* is greater than 4, then n^2 is less than 2^n” can be represented by

$\forall n(P\ (n)\ \rightarrow\ Q(n))$, where the domain consists of all positive integers.

We are assuming that $\forall n(P\ (n)\ \rightarrow\ Q(n))$ is true. Note that $P(100)$ is true because $100 > 4$. It follows by universal modus ponens that
$Q(100)$ is true, namely that $100^2 < 2^{100}$. ■

●Universal Modus Tollens

Another useful combination of a rule of inference from propositional logic and a rule of inference for quantified statements is universal modus tollens. Universal modus tollens combines universal instantiation and modus tollens and can be expressed in the following way:

$$\begin{array}{rl} & \forall x(P(x) \rightarrow Q(x)) \\ & \neg Q(a), \text{ where } a \text{ is a particular element in the domain} \\ \hline \therefore & \neg P(a) \end{array}$$

The verification of universal modus tollens is left as exercise.

Exercise.

Justify the rule of universal modus tollens by showing that the premises $\forall x(P\ (x)\ \rightarrow\ Q(x))$ and $\neg Q(a)$ for a particular element a in the domain, imply $\neg P\ (a)$.

Exercise Set (2.3)

1. Find the argument form for the following argument and determine whether it is valid. Can we conclude that the conclusion is true if the premises are true?

(a) If Socrates is human, then Socrates is mortal.

Socrates is human.

$\therefore$ Socrates is mortal.

(b) If George does not have eight legs, then he is not an insect.

George is an insect.

$\therefore$ George has eight legs.

2. What rule of inference is used in each of these arguments?

(a) Ahmad is a mathematics major. Therefore, Ahmad is either a mathematics major or a computer science major.

(b) Aly is a mathematics major and a computer science major. Therefore, Aly is a mathematics major.

(c) If it is rainy, then the pool will be closed. It is rainy. Therefore, the pool is closed.

(d) If it snows today, the university will close. The university is not closed today. Therefore, it does not snow today.

3. Use rules of inference to show that the hypothesis "Randy works hard" "If Randy works hard, then he is a dull boy", and "If Randy is a dull boy, then he will not get the job" imply the conclusion "Randy will not get the job".

4. What rules of inference are used in this argument?
"No man is an island", "Aly is an island".
Therefore, "Aly is not a man".

5. Show that the argument form with premises
$(p \wedge t) \rightarrow (r \vee s)$, $q \rightarrow (u \wedge t)$, $u \rightarrow p$, and $\neg s$
and conclusion $q \rightarrow r$ is valid by using rules of inference.

6. For each of these arguments, explain which rules of inference are used for each step.

(a) "Sami, a student in this class, knows how to write programs in JAVA. Everyone who knows how to write programs in JAVA can get a high-paying job. Therefore, someone in this class can get a high-paying job"

(b) "Somebody in this class enjoys whale watching. Every person who enjoys whale watching cares about ocean pollution. Therefore, there is a person in this class who cares about ocean pollution".

7. Determine whether the following argument is valid or invalid.

(a)

$$\begin{array}{c} p \to q \vee r \\ p \to \neg q \\ r \to \neg s \\ \text{— — — —} \\ p \to \neg s \end{array}$$

(b)

$$\begin{array}{c} p \to q \\ q \to (p \to r) \\ p \\ \text{— — — —} \\ r \end{array}$$

(c)

$$\begin{array}{c} p \to (q \to r) \\ r \to \neg s \\ \neg u \to s \\ p \wedge q \\ \text{— — — —} \\ u \end{array}$$

(d)

$$\begin{array}{c} p \to q \\ \neg p \to r \\ r \to s \\ \neg q \to s \\ \text{— — — —} \\ q \end{array}$$

(e)

$$\begin{array}{c} p \to q \\ \neg q \vee s \\ q \leftrightarrow s \\ q \to (p \vee \neg s) \\ \text{— — — —} \\ p \leftrightarrow q \end{array}$$

(f)

$$\begin{array}{c} \neg p \to (p \vee r) \\ \neg q \to (\neg q \wedge s) \\ s \to q \vee r \\ \text{— — — —} \\ q \end{array}$$

CHAPTER (III)

METHODS OF PROOF

CHAPTER (III)

METHODS OF PROOF

3.1 Introduction

● Some Terminology

An axiom is a statement that is given to be true.

A rule of inference is a logical rule that is used to deduce one statement from others.

A theorem is a proposition that can be proved using definitions, axioms, other theorems, and rules of inference. Less important theorems sometimes are called **propositions**. (Theorems can also be referred to as **facts** or **results**). A theorem may be the universal quantification of a conditional statement with one or more premises and a conclusion. However, it may be some other type of logical statements. We demonstrate that a theorem is true with a **proof**.

A lemma is a pre-theorem or a result which is needed to prove a theorem.

A corollary is a post-theorem or a result which follows from a theorem (or lemma or another corollary).

A definition is not theorem.

Example of definition: A number n is a perfect square if $n = k^2$ for some integer k.

Definitions are automatically “if and only if” even though they do not say so.

A **proof** is a valid argument that establishes the truth of a theorem. The statements used in a proof can include axioms, premises, if any, of the theorem and previously proven theorems.

A conjecture is a statement that is being proposed to be a true statement, usually on the basis of some partial evidence. When a proof of a conjecture is found, the conjecture becomes a theorem.

3.2 Methods of Proving Theorems

To prove a theorem of the form $\forall x(P(x) \rightarrow Q(x))$, our goal is to show that $P(a) \rightarrow Q(a)$ is true, where a is an arbitrary element of the domain, and then apply universal generalization. In this proof, we need to show that a conditional statement is true. Because of this we now focus on methods that show that conditional statements are true.

♣ Direct Proofs

A direct proof shows that a conditional statement $p \to q$ is true by showing that if p is true, then q must also be true, so that the combination p true and q false never occurs. In a direct proof, we assume that p is true and use axioms, definitions, and previously proven theorems together with rules of inference, to show that q must also be true.

Example.

Direct proof can be used to establish that the sum of two even integers is always even:

(1) x and y are even integers (Hypothesis)

(2) $x = 2a, y = 2b$ for integers a and b (Definition)

(3) $x + y = 2a + 2b = 2(a + b)$ (Algebra)

(4) $x + y$ is even integers (Definition).■

Example.

Between every two distinct rationals, there is a rational.

Proof.

Let $r, s \in \mathbb{Q}$ and $r < s$. Let $t = (r + s)/2$.

Then $t \in \mathbb{Q}$. We must show that $r < t < s$.

Given: $r < s$.

Add r: $2r < r + s$.

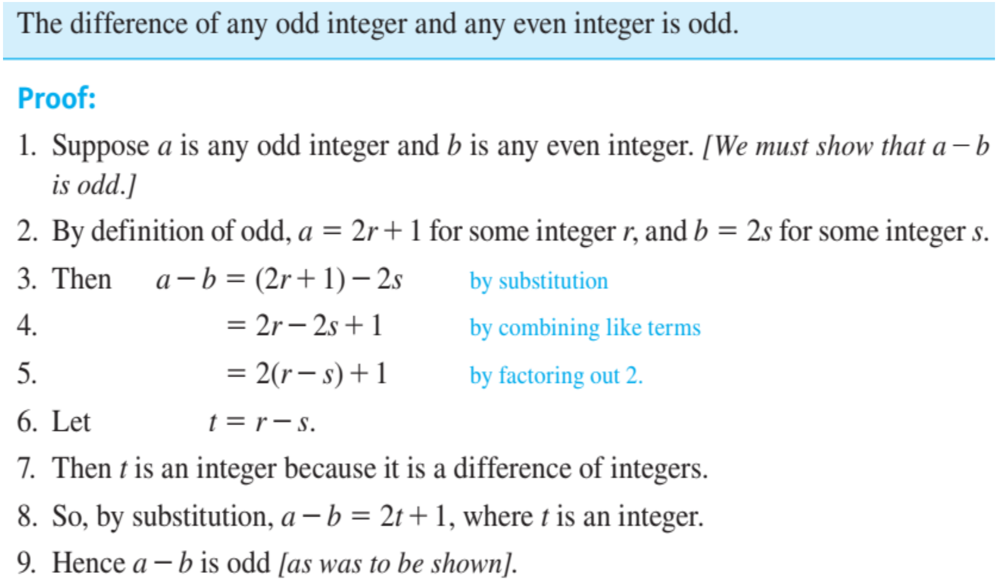
Divide by 2: $r < (r + s)/2 = t$.

Given: $r < s$. Add s: $r + s < 2s$.

Divide by 2: $t = (r + s)/2 < s$.

Therefore $r < t < s$.■

Example.



The difference of any odd integer and any even integer is odd.

Proof:

1. Suppose a is any odd integer and b is any even integer. *[We must show that $a-b$ is odd.]*
2. By definition of odd, $a = 2r+1$ for some integer r, and $b = 2s$ for some integer s.
3. Then $a-b = (2r+1)-2s$ by substitution
4. $= 2r-2s+1$ by combining like terms
5. $= 2(r-s)+1$ by factoring out 2.
6. Let $t = r-s$.
7. Then t is an integer because it is a difference of integers.
8. So, by substitution, $a-b = 2t+1$, where t is an integer.
9. Hence $a-b$ is odd *[as was to be shown]*.

Example.

Give a direct proof of the theorem "If n is an odd integer, then n^2 is odd".

Proof.

(1) n is an odd integer (Hypothesis)

(2) There exists $k \in Z$ such that $n = 2k + 1$

(Definition)

(3) $n^2 = (2k + 1)^2 = 4k^2 + 4k + 1$

$= 2(2k^2 + 2k) + 1$ (Algebra)

(4) n^2 is an odd integer (Definition)

Consequently, we have proved that if n is an odd integer, then n^2 is an odd. ■

Example.

Prove that the statement “The sum of two irrationals is irrational” is false.

Proof.

Counterexample:

Let α be irrational. Then $-\alpha$ is irrational. $\alpha + (-\alpha) = 0$, which is rational.■

Exercise.

Give a direct prove ``If a number is divisible by 6, then it is also divisible by 3".

Definition.

Let *n* be a positive integer. The *n*th *triangle number* T_n is the number $n(n + 1)/2$.

Definition.

Let *n* be a positive integer. The *n*th *perfect square* S_n is the number n^2.

Example.

Give a direct proof to the following statement:

"The sum of two consecutive triangle numbers is a perfect square."

Proof.

Let n be a positive integer.

$$\begin{aligned} T_n + T_{n+1} &= n(n+1)/2 + (n+1)(n+2)/2 \\ &= (n^2 + n + n^2 + 3n + 2)/2 \\ &= (2n^2 + 4n + 2)/2 \\ &= (n+1)^2 \\ &= S_{n+1}. \end{aligned}$$

Therefore, $T_n + T_{n+1} = S_{n+1}$ for all $n \geq 1$.■

Theorem.

If $x, y \in \mathbb{R}$, then $x^2 + y^2 \geq 2xy$.

Incorrect proof:

$x^2 + y^2 \geq 2xy$, $x^2 + y^2 - 2xy \geq 0$.

$(x - y)^2 \geq 0$, which is known to be true.

What is wrong?■

Definition.

Let *x* be a real number.

(a) The ***floor*** of *x* denoted $\lfloor x \rfloor$, is the integer *n* such that $n \leq x < n + 1$. If *x* is an integer, then $\lfloor x \rfloor = x$. If *x* is not an integer, then $\lfloor x \rfloor$ is the first integer such that $\lfloor x \rfloor < x$.

(b) The ***ceiling*** of *x* denoted $\lceil x \rceil$, is the integer *n* such that $n - 1 < x \leq n$. If *x* is an integer, then $\lceil x \rceil = x$.

If *x* is not an integer, then $\lceil x \rceil$ is the first integer such that $\lceil x \rceil > x$.◄

Theorem.

Let *x* and *y* be real numbers. Then

$\lfloor x \rfloor + \lfloor y \rfloor \leq x + y < \lfloor x + y \rfloor + 1.$

Direct Proof.

(1st inequality): By definition, $\lfloor x \rfloor \leq x$ and $\lfloor y \rfloor \leq y$.

Therefore, $\lfloor x \rfloor + \lfloor y \rfloor \leq x + y$.

(2nd inequality): By definition, $x + y < \lfloor x + y \rfloor + 1$.◄

● Proof by Contraposition (indirect proof)

Direct proofs begin with the premises, continue with a sequence of deductions, and end with the conclusion. However, we will see that attempts at direct proofs often reach dead ends. We need other methods of proving theorems of the form $\forall x(P(x) \rightarrow Q(x))$. Proofs of theorems of this type that are not direct proofs, are called **indirect proofs**. An extremely useful type of indirect proof is known as **proof by contraposition**. Proofs by contraposition make use of the fact that the statement $p \rightarrow q$ is equivalent to its contrapositive, $\neg q \longrightarrow \neg p$. This means that the conditional statement $p \rightarrow q$ can be proved by showing that its contrapositive, $\neg q \longrightarrow \neg p$, is true.

Example.

Prove that if n is an integer and $3n + 2$ is odd, then n is odd.

Proof.

We first attempt a direct proof.

(1) 3n + 2 is odd (Hypothesis)

(2) $3n + 2 = 2k + 1$ for $k \in Z$ (Definition)

(3) $3n + 1 = 2k$ (Algebra)

We see that $3n + 1 = 2k$ but there does not seen to be any direct way to conclude that n is odd.

Because our attempt at a direct proof failed, we next try a proof by contraposition.

The contrapositive of "If $3n + 2$ is odd, then n is odd" is "If n is even, then $3n + 2$ is even".

(1) n is even (Hypothesis)

(2) $n = 2k\ , k \in Z$ (Definition)

(3) $3n + 2 = 6k + 2$ (Algebra)

From (3), $3n + 2 = 6k + 2$ is even. Then the given statement is true. ■

Example.

Prove that if $n = ab\ , a, b \in Z^+$, then $a \leq \sqrt{n}$ or $b \leq \sqrt{n}$

Proof.

Because there is no obvious way of showing that $a \leq \sqrt{n}$ or $b \leq \sqrt{n}$ directly from the equation $n = ab$, where a and b are positive integers, we attempt a proof by contraposition.

(1) $a > \sqrt{n}$ and $b > \sqrt{n}$ (Hypothesis)

(2) $ab > \sqrt{n}\,\sqrt{n} = n$ (Algebra)

(3) $ab \neq n$ (Algebra)

Therefore the negation of the conclusion implies that the hypothesis is false. Then the original conditional statement is true. ■

● Proof by Contradiction

Because the statement $r \wedge \neg r$ is a contradiction whenever r is a proposition, we can prove that p is true if we can show that $\neg p \rightarrow (r \wedge \neg r)$ is true for some proposition r. Proofs of this type are called **proofs by contradiction**. Because a proof by contradiction does not prove a result directly, it is another type of indirect proof.

Example.

Prove that $\sqrt{2}$ is irrational by giving a proof by contradiction.

Solution.

(1) Suppose that $\sqrt{2}$ is rational. (Hypothesis)

(2) $\sqrt{2} = \frac{a}{b}$, $a \in Z$, $b \in Z^{*}$ and $\gcd(a, b) = 1$.

(Definition and hypothesis)

(3) $2 = \frac{a^2}{b^2}$ (Algebra)

(4) $a^2 = 2b^2$ (Algebra)

(5) a^2 is even (Definition)

(6) a is even (Algebra)

(7) $a = 2c\ , c \in Z$ (Definition)

(8) $b^2 = 2c^2$ from (7) and (4)

(9) b is even (Algebra)

(10) $\gcd(a, b) \neq 1$ (a contradiction)

Hence our hypothesis that $\sqrt{2}$ is rational is false and hence $\sqrt{2}$ is irrational. ◄

● Proofs of Equivalent

To prove a theorem that is a biconditional statement, that is a statement of the form $p \leftrightarrow q$, we show that $p \rightarrow q$ and $q \rightarrow p$ are both true. The validity of this approach is based on the tautology:

$$(p \leftrightarrow q) \equiv (p \rightarrow q) \wedge (q \rightarrow p)$$

Example.

Prove the theorem "If n is a positive integer, then n is odd if and only if n^2 is odd".

Solution.

This theorem has the form "p if and only q", where p is "n is odd" and q is " n^2 is odd. To prove this theorem, we need to show that $p \rightarrow q$ and $q \rightarrow p$ are both true.

Because we have shown before that both $p \rightarrow q$ and $q \rightarrow p$ are true, we have shown that the theorem is true. ◀

Sometimes a theorem states that several propositions are equivalent. Such a theorem states that propositions$p_1, p_2, \ldots, p_n$ are equivalent. This can be written as $p_1 \leftrightarrow p_2 \leftrightarrow \ldots \leftrightarrow p_n$which states that all n propositions have the same truth values and consequently, that for all i and j with $1 \leq i \leq n$and$1 \leq j \leq n, p_i$and$p_j$are equivalent. One way to prove these mutually equivalent is to use the tautology

$$[p_1 \leftrightarrow \ldots \leftrightarrow p_n] \\ \leftrightarrow [(p_1 \rightarrow p_2) \wedge (p_2 \rightarrow p_3) \wedge \ldots \wedge (p_n \rightarrow p_1)]$$

This shows that if the implications

$p_1 \rightarrow p_2, p_2 \rightarrow p_3, \ldots, p_n \rightarrow p_1$

can be shown to be true, then the proposition $p_1, p_2, \ldots, p_n$are all equivalent.

Example.

Show that these statements are equivalent:

p_1: n is an even integer.

p_2: $n -$ 1 is an odd integer.

p_3: n^2 is an even integer.

Solution.

We use a direct proof to show $p_1 \to p_2$. Suppose that n is even. Then $n = 2k$ for some integer k. Consequently, $n - 1 = 2k - 1 = 2(k - 1) + 1$. This means that $n - 1$ is odd since it is of the form $2m + 1$, where m is the integer $k - 1$.

We also use a direct proof to show $p_2 \to p_3$.

Now, suppose that $n - 1$ is odd.

Then we have $n - 1 = 2k + 1$ for some integer k.

Hence

$$n = 2k + 2.$$

Therefore

$$n^2 = (2k + 2)^2 = 4k^2 + 8k + 4 = 2(2k^2 + 4k + 2).$$

This means that n^2 is even.

To prove $p_3 \to p_1$, we use an indirect proof. That is, we prove that if n is not even, then n^2 is not even. This is the same as proving that if n is odd, then n^2 is odd, which we leave it as exercise. ■

● Disproving Universal Statements

Construct an instance for which the statement $\forall x P(x)$ is false. Also called **Proof by Counterexample.**

Example.

Disprove the statement: If a function is continuous at a point, then it is differentiable at that point.

(Dis)proof:

Let $f(x) = |x|$ and consider the point $x = 0$.

$f(x)$ is continuous at 0. $f(x)$ is not differentiable at 0.◀

Example.

Show that the statement "Every positive integer is the sum of the squares of three integers" is false.

Solution.

To look for a counterexample, we try to write successive positive integers as a sum of three squares. We find that $1 = 0^2 + 0^2 + 1^2$, $2 = 0^2 + 1^2 + 1^2$, $3 = 1^2 + 1^2 +$

1^2 , $4 = 0^2 + 0^2 + 2^2$, $5 = 0^2 + 1^2 + 2^2$, $6 = 1^2 + 1^2 + 2^2$ but we cannot find a way to write 7 as the sum of three squares. It follows that 7 is a counterexample. ■

● Proof By Cases

A proof by cases must cover all possible cases that arise in a theorem.

Example.

Prove that if n is an integer then $n^2 \geq n$.

Solution.

We can prove that $n^2 \geq n$ for every integer by considering three cases,

Case (i). $n = 0$, because $0^2 = 0$, we see that $0^2 \geq 0$. It follows that $n^2 \geq n$ is true in this case.

Case (ii). $n \geq 1$, when we multiply both sides of the inequality $n \geq 1$ by the positive integer n, we obtain $n \cdot n \geq n \cdot 1$. This implies that $n^2 \geq n$ for $n \geq 1$.

Case (iii). $n \leq -1$. Thus, $n^2 \geq 0$. It follows that $n^2 \geq n$.

Because the inequality $n^2 \geq n$ holds in all three cases, we can conclude that if n is an integer, then $n^2 \geq n$. ■

Example.

Use a proof by cases to show that $|xy| = |x||y|$, where x and y are real numbers.

Solution.

We have four cases

Case (i).

We have $xy \geq 0$ when $x \geq 0$ and $y \geq 0$, so that

$|xy| = xy = |x||y|$.

Case (ii).

Note that if $x \geq 0$ and $y < 0$, then $xy \leq 0$ so that

$|xy| = -xy = x(-y) = |x||y|$.

Case (iii).

Note that if $x < 0$ and $y \geq 0$, then $xy \leq 0$ so that

$|xy| = -xy = (-x)y = |x||y|$.

Case (iv).

Note that when $x < 0$ and $y < 0$, it follows that $x\,y > 0$. Hence $|xy| = xy = (-x)(-y) = |x||y|$.

This completes the proof. ■

●Existence Proofs

Proofs of existential statements $\exists x P(x)$ are also called *existence proofs*. Two types of existence proofs

(a) **Constructive**: Construct the object (Prove that it has the necessary properties).

(b) **Non-constructive:** Argue indirectly that the object must exist.

Example.

Between any two distinct irrationals there is a rational and an irrational.

Constructive Proof.

Let α and β be irrational numbers with $\alpha < \beta$.

Then $\beta - \alpha > 0$.

Choose an integer n such that $n(\beta - \alpha) > 1$.

Then $\frac{1}{n} < \beta - \alpha$.

Let $m = \lceil n\beta \rceil - 1$.

Then $m < n\beta \leq m + 1$.

Or $m/n < \beta$ and $n\beta - 1 \leq m$.

Then $\alpha < \beta - 1/n = (n\beta - 1)/n \leq m/n$.

Therefore, $\alpha < m/n < \beta$.

Choose an integer k such that $k(\beta - m/n) > \sqrt{2}$.

Divide by k: $\beta - m/n > \sqrt{2}/k$.

Then $\beta > m/n + \sqrt{2}/k$.

Therefore, $\alpha < m/n < m/n + \sqrt{2}/k < \beta$. ■

Example.

The equation $x^2 - 7y^2 = 1$ has a solution in positive integers.

Constructive proof.

Let $x = 8$ and $y = 3$. Then $8^2 - 7 \times 3^2 = 64 - 63 = 1$.■

Example.

There exists $x \in R$ such that $x^5 - 3x + 1$.

Non-constructive proof.

Let $f(x) = x^5 - 3x + 1$.

$f(1) = -1 < 0$ and $f(2) = 27 > 0$.

$f(x)$ is a continuous function. By the Intermediate Value Theorem, there exists $x \in [1, 2]$ such that $f(x) = 0$.

Example.

Show that there exist irrational numbers a and b such that a^b is rational.

Solution.

We know that $\sqrt{2}$ is irrational. Consider the number

$\sqrt{2}^{\sqrt{2}}$.

Case 1.

If it is rational, we have two irrational numbers a and b, namely $a = \sqrt{2}$ and $b = \sqrt{2}$.

Case 2.

If $\sqrt{2}^{\sqrt{2}}$ is irrational, then we can let $a = \sqrt{2}^{\sqrt{2}}$ and $b = \sqrt{2}$ so $a^b = \sqrt{2}^{\sqrt{2}^{\sqrt{2}}} = \sqrt{2}^{\sqrt{2}\sqrt{2}} = \sqrt{2}^2 = 2.$

This proof is an example of a non-constructive existence proof because we have not found irrational numbers a and b such that a^b is rational. Rather, we have shown that either the pair $a = \sqrt{2}$ and $b = \sqrt{2}$ or the pair $a = \sqrt{2}^{\sqrt{2}}$ and $b = \sqrt{2}$ have desired property, but we do not know which of these two pairs works. ■

Exercises Set (3)

1. Use the direct proof to prove that

(a) $(\forall x)(\forall y)(\forall z)(x + z = y + z \to x = y)$.

(b) if x, y are two rational numbers, then $x + y$ is rational.

(c) if for $a, b, c \in \mathbb{Z}$, $a\backslash b$ and $a\backslash c$, then $a\backslash(bx + cy)$, where $x, y \in \mathbb{Z}$.

(d) if $a\backslash b$ and $b\backslash c$, then $a\backslash c$ for $a, b, c \in \mathbb{Z}$.

2. Let x be a positive real number. Then x is irrational iff the two sequences $\lfloor 1 + x \rfloor, \lfloor 2 + 2x \rfloor, \lfloor 3 + 3x \rfloor, \ldots$

and

$\lfloor 1 + 1/x \rfloor, \lfloor 2 + 2/x \rfloor, \lfloor 3 + 3/x \rfloor, \ldots$

together contain every positive integer exactly once.

3. Let x and y be real numbers.

(a) Prove that $\lceil x + y \rceil - 1 < x + y \leq \lceil x \rceil + \lceil y \rceil$.

(b) Is $-\lfloor -x \rfloor = \lceil x \rceil$ true for all real numbers x?

(c) Is $x - 1 < \lfloor x \rfloor \leq x$ true for all real numbers x?

(d) Is $\lfloor 2x \rfloor + \lfloor 2y \rfloor = \lfloor x \rfloor + \lfloor y \rfloor + \lfloor x + y \rfloor$.

4. Use proof by contraposition to prove that

(a) if $n \in Z$ with n^2 is odd, then n is odd.

(b) if there is no integer between 0 and 1 , then there is no integer between n and $n + 1$.

(c) for $x \in \mathbb{Z}$, if $3|x^2$, then $3|x$.

5. Use the proof by contradiction to show that at least four of any 22 days must fall on the same day of the week.

6. Give a proof by contradiction of the theorem "If $3n + 2$ is odd, then n is odd".

7- Prove that the square of an even number is an even number using

(a) a direct proof.

(b) An indirect proof.

(c) a proof by contradiction.

8- Prove that if x and y are real numbers, then

$$\max(x, y) + \min(x, y) = x + y.$$

9- Prove that the sum of two rational numbers is rational.

10- Show that these three statements are equivalent, where a and b are real numbers:

(a) $a < b$;

(b) $a < \frac{a+b}{2}$;

(c) $\frac{a+b}{2} < b$.

11- Show that if a, b and c are real numbers and$a \neq 0$, then there is a unique solution of the equation $a\,x + b = c$.

12- Prove the triangle inequality, which states that if x and y are real numbers, then$|x| + |y| \geq |x + y|$.

13. Show that there is a positive integer that can be written as the sum of cubes of positive integers in two different ways (give constructive proof).

14. Show that the equation $x^2 – 67y^2 = 1$ has a solution in positive integers. (Give constructive proof).

15. Disprove the conjecture (Fermat): All integers of the form $2^{2^n} + 1$ for $n \geq 1$ are primes. (Give counterexample $n = 5$).

16- Determine whether these are valid arguments.

(a) “If $x^2$is irrational, then x is irrational. Therefore, if x is irrational, it follows that x^2 is irrational”.

(b) “If x^2 is irrational, then x is irrational. The number $x = \pi^2$ is irrational. Therefore, the number $x = \pi$ is irrational”.

CHAPTER (IV)

METHMATICAL INDUCTION

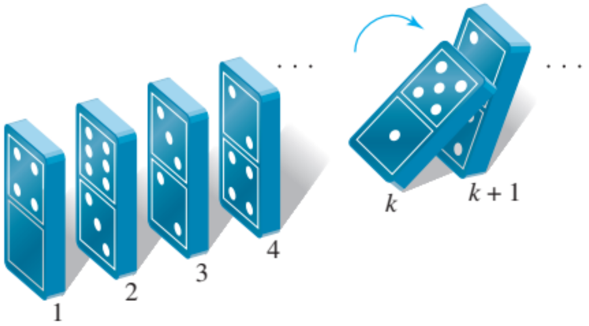
Chapter (IV)

Mathematical Induction

4.1 The Basic Principle

The basic principle of mathematical induction is as follows. To prove that a statement holds for all positive integers n, we first verify that it holds for $n = 1$, and then we prove that if it holds for a certain natural number k, it also holds for $k + 1$.

To visualize the idea of mathematical induction, imagine an infinite collection of dominoes positioned one behind the other in such a way that if any given domino falls backward, it makes the one behind it fall backward also. Then imagine that the first domino falls backward. What happens? Á They all fall down!



If the kth domino falls backward, it pushes the $(k + 1)$st domino backward also.

Theorem 1.

(Principle of Mathematical Induction)

Let $S(n)$ denote a statement involving a variable n. Suppose

(1) $S(1)$ is true ;

(2) if $S(k)$ is true for some positive k, then $S(k+1)$ is also true.

Then $S(n)$ is true for all positive integers n.

Example.

Prove $1+3+5+\cdots+(2n-1)=n^2$ for all natural numbers n.

Solution.

We shall prove the statement using mathematical induction. Clearly, the statement holds when $n=1$ since $1=1^2$. Suppose the statement holds for some positive integer k.

That is, $1+3+5+\cdots+(2k-1)=k^2$.

Consider the case $n=k+1$. By the above assumption (which we shall call the induction hypothesis), we have

$1+3+5+\cdots+[2(k+1)-1]$

$$= [1 + 3 + 5 + \cdots + (2k - 1)] + (2k + 1)$$
$$= k^2 + (2k + 1) \ = (k + 1)^2$$

That is the statement holds for $n = k + 1$ provided that it holds for $n = k$. By the principle of mathematical induction, we conclude that $1 + 3 + 5 + \cdots + (2n - 1) = n^2$ for all natural numbers n. ■

The principle of mathematical induction can be used to prove a wide range of statements involving variables that take discrete values. Some typical examples are shown below.

Example.

Prove that $23^n - 1$ is divisible by 11 for all positive integers n.

Solution.

Clearly $23^1 - 1 = 22$ is divisible by 11. Suppose $11 \big| 23^k$ for some positive integer k.

For the case $n = k + 1$, we have

$$23^{k+1} - 1 = 23.23^k - 1 = 11.2.23^k + (23^k - 1)$$

which is also divisible by 11. It follows that $23^n - 1$ is divisible by 11 for all positive integers n. ■

Example.

Let $x > -1$ be a real number.

Prove that $(1+x)^n \geq 1+nx$ for all natural numbers n.

Solution.

The inequality holds for $n = 1$ since $(1+x)^1 = 1+1x$.

Let $(1+x)^k \geq 1+kx$ for some positive integer k. For the case $n = k+1$,

$$\begin{aligned}(1+x)^{k+1} &= (1+x)^k(1+x) \geq (1+kx)(1+x)\\ &= 1+(k+1)x+kx^2 \geq 1+(k+1)x.\end{aligned}$$

Hence, if the inequality holds for the case $n = k$, it also holds for the case $n = k+1$. It follows that $(1+x)^n \geq 1+nx$ for all natural numbers n. ■

4.2 Variations of the Basic principle

There are many variations to the principle of mathematical induction.

Theorem 2. (Principle of Mathematical Induction, Variation 1)

Let $S(n)$ denote a statement involving a variable n. Suppose

(1) $S(k_0)$ is true for some positive integer k_0 ;

(2) if $S(k)$ is true for some positive integer $k \geq k_0$, then

$S(k+1)$ is also true.

Then $S(n)$ is true for all positive integers $n \geq k_0$.

In some cases a statement involving a variable n holds when n is 'Large enough', but does not hold when, say, $n = 1$. In this case Theorem 1 does not apply, but the above variation allows us to prove the statement.

Example.

Prove that $2^n > n^2$ for all natural numbers $n \geq 5$.

Solution.

First, we check that $2^5 = 32 > 25 = 5^2$, so the inequality holds for $n = 5$.

Suppose $2^k > k^2$ for some integer $k \geq 5$.

Then $2^{k+1} = 2.2^k > 2k^2 > (k+1)^2$.

The last inequality holds since $2k^2 - (k+1)^2 = (k-1)^2 - 2 > 0$ whenever $k \geq 5$.

Hence, if the inequality holds for $n = k$, it also holds for $n = k + 1$. By Theorem 2, $2^n > n^2$ for all natural numbers $n \geq 5$. ■

Sometimes a sequence may be defined recursively, and a term may depend on some previous terms. In particular, it may depend on more than one previous terms. In this case Theorem 1 does not apply because assuming $S(k)$ holds for a single k is not sufficient. We need the following.

Theorem 3. (Principle of Mathematical Induction, Variation 2)

Let $S(n)$ denote a statement involving a variable n. Suppose

(1) $S(1)$ and $S(2)$ are true;

(2) if $S(k)$ and $S(k+1)$ are true for some positive integer k, then $S(k+2)$ is also true.

Then $S(n)$ is true for all positive integers n.

Of course there is no need to restrict ourselves only to ' two levels'. Moreover, in the spirit of Theorem 2, there is no need to start from $n = 1$. We leave the formulation as an exercise.

Example.

Let $\{a_n\}$ be a sequence of natural numbers such that $a_1 = 5, a_2 = 13$ and $a_{n+2} = 5a_{n+1} - 6a_n$ for all natural numbers n. Prove that $a_n = 2^n + 3^n$ for all natural numbers n.

Solution.

We have that $a_1 = 5 = 2^1 + 3^1$ and $a_2 = 13 = 2^2 + 3^2$. Suppose $a_k = 2^k + 3^k$ and $a_{k+1} = 2^{k+1} + 3^{k+1}$ for some natural number k.

Then

$$\begin{aligned} a_{k+2} &= 5a_{k+1} - 6a_k \\ &= 5(2^{k+1} + 3^{k+1}) - 6(2^k + 3^k) \\ &= 4.2^k + 9.3^k = 2^{k+2} + 3^{k+2} \end{aligned}$$

Hence, if the formula holds for $n = k$ and $n = k + 1$, it also holds for $n = k + 2$. By Theorem 3 we have $a_n = 2^n + 3^n$ for some natural number n. ■

Sometimes to prove a statement we need to consider the odd cases and even cases separately. To combine them nicely into one single case, we need the following.

Theorem 4. (Principle of Mathematical Induction, Variation 3)

Let $S(n)$ denote a statement involving a variable n. Suppose

(1) $S(1)$ and $S(2)$ are true;

(2) if $S(k)$ is true for some positive integer k, then

$S(k + 2)$ is also true.

Then $S(n)$ is true for all positive integers n.

Although Theorem 2 and Theorem 3 look similar, their nature is quite different. Again there is no need to restrict ourselves to considering only two initial cases, but we do not bother to go into the details.

Example.

Prove that for all natural numbers n, there exist distinct integers x, y, z for which

$$x^2 + y^2 + z^2 = 14^n.$$

Solution. For $n = 1$ and $n = 2$, such integers exist as $1^2 + 2^2 + 3^2 = 14$ and $4^2 + 6^2 + 12^2 = 14^2$.

Suppose for $n = k$ (where k is positive integer), such integers exist, i.e. $x_\circ^2 + y_\circ^2 + z_\circ^2 = 14^k$ for some distinct integers $x_\circ, y_\circ, z_\circ$.

Then for $n = k + 2$, such integers also exist because

$$(14x_\circ)^2 + (14y_\circ)^2 + (14z_\circ)^2 = 14^{k+2}.$$

By Theorem 4, the result follows. ■

In Theorem 2, we remarked that sometimes assumption of $S(k)$ for a single k may not be sufficient, so we may need to assume the statement holds for two values (and accordingly we need to verify two initial cases).

We also remarked that there is no need to restrict ourselves to only two values; we could generalize to any finite number of cases.

The following variation gives a further generalization of this, assuming all cases from 1 to k.

Theorem 5.

(Principle of Mathematical Induction, Variation 4)

Let $S(n)$ denote a statement involving a variable n. Suppose

(1) $S(1)$ is true;

(2) if for some positive integer k, $S(1), S(2), \dots, S(k)$ are all true, then $S(k+1)$ is also true.

Then $S(n)$ is true for all positive integers n.

Example.

Let $a_1, a_2, \dots$ be a sequence of real numbers satisfying $a_{i+j} \leq a_i + a_j$ for all $i, j = 1, 2, \dots$

Prove that

$$a_1 + \frac{a_2}{2} + \frac{a_3}{3} + \cdots + \frac{a_n}{n} \geq a_n$$

for each positive integer n.

Solution.

Clearly, the inequality holds for $n = 1$.

Suppose the inequality holds for $n = 1,2,\dots,k$ for some positive integer k.

Then by adding the inequalities

$$a_1 \geq a_1$$

$$a_1 + \frac{a_2}{2} \geq a_2$$

$$\vdots$$

$$a_1 + \frac{a_2}{2} + \cdots + \frac{a_k}{k} \geq a_k$$

We get

$$ka_1 + (k-1)\frac{a_2}{2} + \cdots + \frac{a_k}{k} \geq a_1 + a_2 + \ldots + a_k$$

$i, e.,$

$$(k+1)\left(a_1 + \frac{a_2}{2} + \cdots + \frac{a_k}{k}\right) \geq 2(a_1 + a_2 + \cdots + a_k)$$

$$= (a_1 + a_k) + (a_2 + a_{k-1}) + \cdots + (a_k + a_1)$$

$$\geq ka_{k+1}.$$

It follows that

$$(k+1)\left(a_1 + \frac{a_2}{2} + \cdots + \frac{a_k}{k} + \frac{a_{k+1}}{k+1}\right) \geq (k+1)a_{k+1}.$$

Hence

$$a_1 + \frac{a_2}{2} + \cdots + \frac{a_{k+1}}{k+1} \geq a_{k+1},$$

i. e., the inequality holds for $n = k + 1$.

By Theorem 6, the result follows. ■

Finally, we introduce a special variation, commonly known as backward induction.

Theorem 6. (Backward Induction)

Let $S(n)$ denote a statement involving a variable n. Suppose

(1) $S(n)$ is true for infinitely many natural numbers n ;

(2) if $S(k)$ is true for some positive integer $k > 1$, then $S(k-1)$ is also true.

Then $S(n)$ is true for all positive integers n.

The most typical example backward induction is used is perhaps in the proof of the **AM-GM** inequality, as shown in the example below.

Example. (AM-GM Inequality)

Prove that for positive integers $a_1, a_2, \ldots, a_n$,

$$\frac{a_1 + a_2 + \cdots + a_n}{n} \geq \sqrt[n]{a_1 a_2 \ldots a_n}.$$

In other words, the **arithmetic mean** (AM) is always greater than or equal to the **geometric mean** (GM).

Solution.

From $(\sqrt{a_1}-\sqrt{a_2})^2 \geq 0$, we obtain $\frac{a_1+a_2}{2} \geq \sqrt{a_1 a_2}\,, i, e.$ the inequality holds for $n = 2$. Suppose the inequality holds when $n = k$ for some positive integer k. Consider the case $n = 2k$. Using the case $n = 2$ and the induction hypothesis, we have

$$\frac{a_1+a_2+\cdots+a_{2k}}{2k} = \frac{1}{k}\left(\frac{a_1+a_2}{2}+\frac{a_3+a_4}{2}+\cdots+\frac{a_{2k-1}+a_{2k}}{2}\right)$$

$$\geq \frac{\sqrt{a_1a_2}+\sqrt{a_3a_4}+\cdots+\sqrt{a_{2k-1}a_{2k}}}{k}$$

$$\geq \sqrt[k]{\sqrt{a_1a_2}.\sqrt{a_3a_4}\ldots\sqrt{a_{2k-1}a_{2k}}}$$

$$\geq \sqrt[2k]{a_1a_2\ldots a_{2k}}$$

$i, e.$ the inequality also holds for $n = 2k$. By Theorem 1, the inequality holds for all positive powers of 2. In other words, condition (1) in Theorem 6 is satisfied. Again, we suppose the inequality holds when $n = k$ for some positive integer $k, i.e.$,

$$\frac{a_1+a_2+\cdots+a_k}{k} \geq \sqrt[k]{a_1a_2\ldots a_k}.$$

Applying the substitution $a_k = \frac{a_1+a_2+\cdots+a_{k-1}}{k-1}$ and simplifying (the details of which are left as an exercise), we get $\frac{a_1+a_2+\cdots+a_{k-1}}{k-1} \geq \sqrt[k-1]{a_1 + a_2 + \cdots + a_{k-1}}$ $i, e.$ the inequality also holds when $n = k - 1.$ By Theorem 6, the inequality is proved. ■

♣ 4.3 Miscellaneous Examples

Most of the examples we have seen deal with algebraic (in) equalities and problems in number theory. One should not be misled to think that these are the only areas in which the method of mathematical induction applies. In fact, the method is powerful that it is useful in almost every branch of mathematics. In this section we shall see some miscellaneous examples.

Example.

Prove that

$$\sin\theta + \sin 2\theta + \cdots + \sin n\theta = \sin\frac{(n+1)\theta}{2}\sin\frac{n\theta}{2}\csc\frac{\theta}{2}$$

for all positive integers n.

Solution.

When $n = 1$, the right hand side is:

$$\sin\theta \sin\frac{\theta}{2}\csc\frac{\theta}{2} = \sin\theta.$$

So the formula holds for $n = 1$.

Suppose the formula holds for $n = k, i.e.$

$\sin\theta + \sin 2\theta + \cdots + \sin k\theta = \sin\frac{(k+1)\theta}{2}\sin\frac{k\theta}{2}\csc\frac{\theta}{2}$

Consider the case $n = k + 1$.

By the induction hypothesis,

$\sin\theta + \sin 2\theta + \cdots + \sin k\theta + \sin(k+1)\theta$

$= \sin\frac{(k+1)\theta}{2}\sin\frac{k\theta}{2}\csc\frac{\theta}{2} + \sin(\mathrm{k}+1)\theta$

$= \sin\frac{(k+1)\theta}{2}\sin\frac{k\theta}{2}\csc\frac{\theta}{2} + 2\sin\frac{(k+1)\theta}{2}\cos\frac{(k+1)\theta}{2}$

$= \sin\frac{(k+1)\theta}{2}\csc\frac{\theta}{2}\left[\sin\frac{k\theta}{2} + 2\sin\frac{\theta}{2}\cos\frac{(k+1)\theta}{2}\right]$

$= \sin\frac{(k+1)\theta}{2}\csc\frac{\theta}{2}\left[\sin\frac{k\theta}{2} + \sin\left(\frac{\theta}{2} + \frac{(k+1)\theta}{2}\right) + \right.$

$\left.\sin\left(\frac{\theta}{2} - \frac{(k+1)\theta}{2}\right)\right]$

$= \sin\frac{(k+1)\theta}{2}\csc\frac{\theta}{2}\left[\sin\frac{k\theta}{2} + \sin\frac{(k+2)\theta}{2} - \sin\frac{k\theta}{2}\right]$

$= \sin\frac{[(k+1)+1]\theta}{2}\sin\frac{(k+1)\theta}{2}\csc\frac{\theta}{2}$

By the principle of mathematical induction, the formula holds for all positive integers n. ■

Example.

Prove that $\left(3+\sqrt{5}\right)^n+\left(3-\sqrt{5}\right)^n$is an even integer for all natural numbers n.

Solution.

Write $f(n)=\alpha^n+\beta^n$ where $\alpha=3+\sqrt{5}$ and $\beta=3-\sqrt{5}$.

It is straightforward to check that $f(1)=6$ and $f(2)=28$ are even integers. Suppose $f(k)$ and $f(k+1)$ are both even integers for some positive integer k. Consider the case $n=k+2$. Note that α and β are roots of the equation $x^2-6x+4=0$.

So $\alpha^2=6\alpha-4$ and $\beta^2=6\beta-4$, and thus

$$\begin{aligned} f(k+2) &= \alpha^{k+2}+\beta^{k+2} \\ &= \alpha^k(6\alpha-4)+\beta^k(6\beta-4) \\ &= 6(\alpha^{k+1}+\beta^{k+1})-4(\alpha^k+\beta^k) \\ &= 6f(k+1)-4f(k) \end{aligned}$$

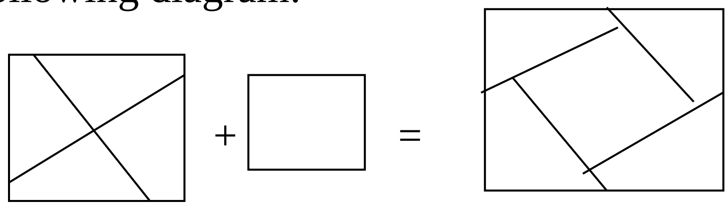
It follows that $f(k+2)$ must also be an even integer.

By mathematical induction, we conclude that $f(n)$ is an even integer for all natural numbers n. ■

Example.

Prove that, given two or more squares, one can always cut them (using only compasses, straight edge and scissors) and reform them into a large square.

Solution. In the case of two squares, we resort to the following diagram:



We leave it to the reader to work how the dotted lines are to be drawn and to verify that such constructions are indeed possible using compasses and straight edge.

Suppose the statement is true in the case of k squares. Then, in the case of $k+1$ squares, we can cut k of the squares to form a large square, according to the induction hypothesis. This leaves only two squares, but we have shown that two squares can be cut to form one large square. By the principle of mathematical induction, the statement is proved. ■

Example.

In a party there are $2n$ participants, where n is a natural number. Some participants shake hands with other participants. It is known that there do not exist three participants who have shaken hands with each other. Prove that the total number of handshakes is not more than n^2 .

Solution.

When $n = 1$, the number of handshakes is at most $1 = 1^2$.

Suppose that with 2k people, the total number of handshakes is at most k^2 under the given condition.

Consider the case $n = k + 1, i.e. 2k + 2$ people.

Pick two people who have shaken hands with each other (if no such people exist, then the total number of handshake would be zero), say A and B. Under the induction hypothesis, there are at most k^2 handshakes among the other $2k$ people.

Now by the given condition, none of these $2k$ people have shaken hands with both A and B. So these $2k$ people have at most $2k$ handshakes with A and B.

Taking the handshake between A and B into account, the total number of handshakes is at most $k^2 + 2k + 1 = (k+1)^2$. By the principle of mathematical induction, the result follows. ■

4.4 Higher Dimensional Induction

Theorem 7. (Two- Dimensional Induction, Version 1)

Let $S(m,n)$ denote a statement involving two variables m and n. Suppose

(1) $S(1,1)$ is true;

(2) if $S(k,1)$ is true for some positive integer k, then $S(k+1,1)$ is also true.

(3) if $S(h,k)$ holds for some positive integer h and k, then $S(h,k+1)$ is also true.

Then $S(m,n)$ is true for all positive integers m,n.

Theorem 7 can be easily understood. The first two conditions together imply (by Theorem 1) that $S(m,1)$ is true for all positive integers m. Thus, fixing m, this together with condition (3) imply (by Theorem 1 again) that $S(m,n)$ holds for all positive integers n. As a result,

$S(m,n)$ holds for all positive integers m and n, as we desire.

Example.

Let f be a function of two variables, with $f(1,1) = 2$,

$$\begin{cases} f(m+1,n) = f(m,n) + 2(m+n) \\ f(m,n+1) = f(m,n) + 2(m+n-1) \end{cases}$$

For all natural numbers m and n. Prove that

$$f(m,n) = (m+n)^2 - (m+n) - 2n + 2$$

For all positive integers m and n.

Solution.

We first check that $f(1,1) = 2 = (1+1)^2 - (1+1) - 2(1) + 2$. Suppose $f(k,1) = (k+1)^2 - (k+1) - 2(1) + 2 = k^2 + k$ for some positive integers k. Then

$$\begin{aligned} f(k+1,1) &= f(k,1) + 2(k+1) \\ &= (k^2+k) + (2k+2) \\ &= [(k+1)+1]^2 - [(k+1)+1] - 2(1) + 2. \end{aligned}$$

Thus conditions (1) and (2) in Theorem 7 are satisfied.

Suppose $f(h,k) = (h+k)^2 - (h+k) - 2k + 2$ for some positive integers h and k. Then

$$\begin{aligned} f(h,k+1) &= f(h,k) + 2(h+k-1) \\ &= (h+k)^2 - (h+k) - 2k + 2 + 2(h+k) - 2 \end{aligned}$$

$$= (h + k + 1)^2 - (h + k + 1) - 2(k + 1) + 2$$

Thus condition (3) in Theorem 7 is also satisfied.

It follows that $f(m, n) = (m + n)^2 - (m + n) - 2n + 2$ for all positive integers m and n. ■

Theorem 7 is essentially applying Theorem 1 twice. The following alternative version of the principle of two-dimensional induction in some sense reduces a two-dimensional problem into one dimension.

Theorem 8. (Two- Dimensional Induction, Version 2)

Let $S(m, n)$ denote a statement involving two variables m and n. Suppose

(1) $S(1,1)$ is true;

(2) if for some positive integer $k > 1$, $S(m, n)$ is true whenever $m + n = k$, then $S(m, n)$ is true whenever $m + n = k + 1$.

Then $S(m, n)$ is true for all positive integers m, n.

Example.

For natural numbers p and q, the Ramsey number $R(p, q)$ is defined as smallest integer n so that among any n people, there exist p of them who know each other, or there exist q of them who don't know each other. (We assume that if A knows B, then B knows A, and vice versa.) It is known that $R(p,1) = R(1,q) = 1$

and $R(p+1,q+1) \leq R(p,q+1) + R(p+1,q)$

For all natural numbers p and q. Deduce for all natural numbers p,q that

$$R(p,q) \leq C_{p-1}^{p+q-2}.$$

Solution.

First, we check that $R(1,1) = 1 = C_{1-1}^{1+1-2}$.

Assume that the desired inequality holds for all p,q with $p+q=k$, where k is a positive integer.

Now consider $R(p,q)$ with $+q = k+1$.

If either $p=1$ or $q=1$, the desired inequality follows immediately.

If not, then noting that $(p-1)+q = p+(q-1) = k$, the inductive hypothesis gives

$$R(p,q) \leq R(p-1,q) + R(p,q-1)$$
$$\leq C_{p-2}^{p+q-3} + C_{p-1}^{p+q-3} = C_{p-1}^{p+q-2}.$$

In other words, the desired inequality holds whenever $p+q=k+1$. By Theorem 8 the result follows.■

Exercise Set (4)

1- Prove by mathematical induction that the following statements hold for all positive integers n.

(a) $1^2 + 2^2 + \cdots + n^2 = \frac{1}{6}n(n+1)(2n+1)$;

(b) $1^2 \times 2 + 2^2 \times 3 + \cdots + n^2(n+1) = \frac{n(n+1)(n+2)(3n+1)}{12}$;

(c) $4007^n - 1$ is divisible by 2003;

(d) $2002^{n+2} + 2003^{2n+1}$ is divisible by 4005;

(e) $n^2 > n + 1$;

(f) $\frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{2^n} \geq \frac{n}{2}$;

(g) $1 \times 1! + 2 \times 2! + \cdots + n \times n! = (n+1)! - 1$;

(h) $\cos\theta + \cos 2\theta + \ldots + \cos n\,\theta = \sin\frac{(n+1)\theta}{2}\cos\frac{n\theta}{2}\csc\frac{\theta}{2} - 1.$

2- To apply the principle of mathematical induction we need to verify two conditions, namely, the statement holds for $n = 1$, and that if statement holds for $n = k$ it also holds for $n = k + 1$. Can you think of a (wrong) statement in which the second condition is satisfied but the first one is not? That is, can you construct a statement $S(n)$ such that if $S(k)$ true, then $S(k+1)$ must be true, yet $S(1)$ is not true?

3- Prove that for all natural numbers n,

$1+\frac{1}{2^2}+\frac{1}{3^2}+\cdots+\frac{1}{n^2}\leq 2-\frac{1}{n}$.

What is the significance of the above result on the convergence of the series $\sum n^{-2}$?

4- The Lucas sequence $1,\ 3,\ 4,\ 7,\ 11,\ 18,\ 29,$ .... is defined by $a_1=1,\ a_2=3,\quad a_n=a_{n-1}+a_{n-2}$ for $n\geq 3$.

Prove that $a_n<(1.75)^n$ for all positive integers n.

5- From a pack of 52 playing cards one extracts the 26 red cards and pairs them up randomly. The back sides of each pair of cards are then glued together, resulting in 13 cards with both sides being 'the front'. Prove that it is always possible to flip the cards so that the 13 sides facing upward are $A, 2, 3, \ldots 10, J,\ Q,\ K$.

6- The Fibonacci sequence is defined as $x_0=0, x_1=1$ and $x_{n+2}=x_{n+1}+x_n$ for all non-negative integers n. Prove that

(a) $x_m=x_{r+1}x_{m-r}+x_rx_{m-r-1}$ for all integers $m\geq 1$ and $0\leq r\leq m-1$;

(b) x_d divides x_{kd} for all positive integers d and k.

CHAPTER (V)

ELEMENTARY NUMBER THEORY

Chapter (V)

Elementary Number Theory

5.1 The Ring of Integers

Elementary number theory is largely about the **ring of integers**, denoted by the symbol $\mathbb{Z}$. The integers are an example of an algebraic structure called an **integral domain**. This means that $\mathbb{Z}$ satisfies the following axioms:

(a) $\mathbb{Z}$ has operations '+' (addition) and '·' (multiplication). It is **closed** under these operations, in that if $m, n \in \mathbb{Z}$, then $m + n \in \mathbb{Z}$ and $m \cdot n \in \mathbb{Z}$.

(b) Addition is **associative**: If $m, n, p \in \mathbb{Z}$, then

$$m + (n + p) = (m + n) + p$$

(c) There is an **additive identity** $0 \in \mathbb{Z}$: For all $n \in \mathbb{Z}$,

$$n + 0 = n \text{ and } 0 + n = n.$$

(d) Every element has an **additive inverse**: If $n \in \mathbb{Z}$, there is an element $-n \in \mathbb{Z}$such that

$$n + (-n) = 0 \text{ and } (-n) + n = 0.$$

(e) Addition is **commutative**: If $m, n \in \mathbb{Z}$, then

$$m + n = n + m.$$

(f) Multiplication is **associative**:

If $m, n, p \in \mathbb{Z}$, then $m \cdot (n \cdot p) = (m \cdot n) \cdot p$.

(g) There is an **multiplicative identity**$1 \in \mathbb{Z}$:

For all $n \in \mathbb{Z}$,$n \cdot 1 = n$ and $1 \cdot n = n$.

(h) Multiplication is **commutative**:

If $m, n \in \mathbb{Z}$ then $m \cdot n = n \cdot m$.

(i) The **Distributive Laws** hold:

If $m, n, p \in \mathbb{Z}$, then $m \cdot (n + p) = m \cdot n + m \cdot p$ and $(m + n) \cdot p = m \cdot p + n \cdot p$.

(j) There are **no zero divisors**:

If $m, n \in \mathbb{Z}$ and $m \cdot n = 0$, then either $m = 0$ or $n = 0$.

Remarks.

(a) As usual, we'll often abbreviate $m \cdot n$ to mn.

(b) The last axiom is equivalent to the **Cancellation Property**: If $a, b, c \in \mathbb{Z}, a \neq 0$, and $ab = ac$, then $b = c$.

Example.

If $n \in \mathbb{Z}$, prove that $0 \cdot n = 0$.

Solution.

$0 \cdot n = (0 + 0) \cdot n$ (Additive identity)

$= 0 \cdot n + 0 \cdot n$ (Distributive Law)

Adding $-(0 \cdot n)$ to both sides, we get

$$-(0 \cdot n) + 0 \cdot n = -(0 \cdot n) + (0 \cdot n + 0 \cdot n)$$

By associativity for addition,

$$-(0 \cdot n) + 0 \cdot n = (-(0 \cdot n) + 0 \cdot n) + 0 \cdot n.$$

Then using the fact that – $(0 \cdot n)$ and $0 \cdot n$ are additive inverses, we get

$$0 = 0 + 0 \cdot n.$$

Finally, 0 is the additive identity, so

$$0 \ = \ 0 \cdot \ n.$$◀

Example.

If $n \in \mathbb{Z}$, prove that $-n = (-1) \cdot n$.

Solution.

In other words, the equation says that the additive inverse of n (namely $-n$) is equal to $(-1) \cdot n$.

What is the additive inverse of n?

It is the number which gives 0 when added to n.

Therefore, we should add $(-1) \cdot n$ and see if I get 0:

$$\begin{aligned}(-1) \cdot n + n &= (-1) \cdot n + 1 \cdot n \text{ (Multiplicative identity)}\\ &= (-1 + 1) \cdot n \quad \text{(Distributive Law)}\\ &= 0 \cdot n \text{ (Additive inverse)}\\ &= 0. \quad \text{(Preceding result)}\end{aligned}$$

This proves that $-n = (-1) \cdot n$.◀

● The integers are **ordered** --- there is a notion of greater than (or less than). Specifically, for $m, n \in \mathbb{Z}, m > n$ is defined to mean that $m - n$ is a **positive integer** --- and element of the set $\{1, 2, 3, \ldots\}$.

Of course, $m < n$ is defined to mean $n > m$.

$m \geq n$ and $n \leq m$ have the obvious meanings.

(k) The positive integers are closed under addition and multiplication.

There are two order axioms:

●Trichotomy:

If $n \in \mathbb{Z}$, either$n > 0, n < 0,$ or$n = 0$.

Example.

Prove that if $m > 0, n < 0,$ then $mn < 0$.

Solution.

Since $n < 0,\ 0 - n = -n$ is a positive integer.

$m > 0$ means $m = m - 0$ is a positive integer, so by closure $m \cdot (-n)$ is a positive integer. By a property of integers (which you should try proving from the axioms), $m \cdot (-n) = -(m \cdot n)$. Thus, $-(m \cdot n)$ is a positive integer. So $0 - mn = -(mn)$ is a positive integer, which means that $mn < 0$.◀

●**The Well-Ordering Property** of the integers sounds simple: Every nonempty subset of the positive integers has a smallest element. Your long experience with the integers makes this principle sound obvious. In fact, it is one of the deeper axioms for $\mathbb{Z}$; for example, it can be used to prove the principle of **mathematical induction**, which we have discussed.

Example.

Prove that $\sqrt[3]{2}$ is irrational number.

Solution.

The proof will use the Well-Ordering Property.

We'll give a proof by contradiction. Suppose that $\sqrt[3]{2}$ is a rational number. In that case, we can write $\sqrt[3]{2} = \frac{a}{b}$,

where a and b are positive integers. Now

$\sqrt[3]{2} = \frac{a}{b}$, so $b\sqrt[3]{2} = a$ and $2b^3 = a^3$.

(To complete the proof, we are going to use some divisibility properties of the integers that we haven't proven yet. They're easy to understand and pretty plausible, so this shouldn't be a problem.)

The last equation shows that 2 divides a^3. This is only possible if 2 divides a, so $a = 2c$, for some positive integer c. Plugging this into $2b^3 = a^3$, we get

$$2b^3 = 8c^3, or\ b^3 = 4c^3.$$

Since 2 divides $4c^3$, it follows that 2 divides b^3. As before, this is only possible if 2 divides b, so $b = 2d$ for some positive integer d. Plugging this into $b^3 = 4c^3$.

We get $8d^3 = 4c^3$, or $2d^3 = c^3$.

This equation has the same form as the equation $2b^3 = a^3$, so it's clear that we can continue this procedure indefinitely to get *e* such that $c = 2e$, *f* such that $d = 2f,$ and so on.

However, since $a = 2c$, it follows that $a > c$; since $c = 2e$, we have $c > e,$ so $a > c > e$. Thus, the numbers $a, c, e, \ldots$. comprise a set of positive integers *with no smallest element*, since a given number in the list is always smaller than the one before it. This contradicts Well-Ordering. Therefore, my assumption that $\sqrt[3]{2}$ is a rational number is wrong, and hence $\sqrt[3]{2}$ is irrational.◀

● Finally, we want to mention a function that comes up often in number theory.

Definition.

If x is a real number, then $[x]$ denotes the **greatest integer function** of x (Is it the floor of x?). It is the largest integer less than or equal to x.

Lemma.

If x is a real number, then $[x] + 1 > x \geq [x]$.

Proof.

By definition, $x \geq [x]$. To show that $[x] + 1 > x$, we'll give a **proof by contradiction**. Suppose on the contrary that $[x] + 1 \leq x$. Then $[x] + 1$ is an integer less than or equal to x, which contradicts the fact that $[x]$ is the *largest* integer less than or equal to x. This contradiction implies that $[x] + 1 > x$.◄

Lemma.

If $x, y \in \mathbb{R}$ and $x \geq y$, then $[x] \geq [y]$.

Proof.

Suppose $x \geq y$. We want to show that $[x] \geq [y]$.

Assume on the **contrary** that $[y] > [x]$. Since $[x]$ is the it greatest integer which is less than or equal to x, and since $[y]$ is an integer which is greater than $[x]$, it follows that $[y]$ can't be less than or equal to x. Thus, $[y] > x$. But $x \geq y$. So $[y] > y$, which is a contradiction. Therefore, $[x] \geq [y]$.◀

Example.

Find $[3.2]$, $[117]$ and $[-1.2]$.

Solution.

$[3.2] = 3, [117] = 117, and\ [-1.2] = -2.$

(Notice that $[-1.2]$ is *not* equal to -1).◀

Example.

Let x be a real number and let n be an integer. Prove that

$$[x + n] = [x] + n.$$

Solution.

First, $x \geq [x]$, so $x + n \geq [x] + n$.

Now, $[x] + n$ is an integer less than or equal to $x + n$, so it must be less than or equal to the greatest integer less than or equal to $x + n$--- which is $[x + n]$:

$$[x + n] \geq [x] + n.$$

Next, $x+n \geq [x+n]$. Then $x \geq [x+n]-n$ and $[x+n]-n$ is an integer less than or equal to x. Therefore, it must be less than or equal to the greatest integer less than or equal to x --- which is$[x]$:

$$[x] \geq [x+n]-n.$$

Adding n to both sides gives

$$[x]+n \geq [x+n].$$

Since $[x+n] \geq [x]+n$ and $[x]+n \geq [x+n]$, it follows that $[x]+n=[x+n]$.◀

5.2 Prime Numbers

●Every integer greater than 1 is divisible by at least two integers, because a positive integer is divisible by 1 and by itself. Positive integers that have exactly two different positive integer factors are called primes.

● Euclid showed that there are infinitely many primes.

● The ***Prime Number Theorem*** says that the number of primes less than or equal to a real number *x* is approximately $\frac{x}{\ln x}$.

●The prime numbers are the "building blocks" of the integers. We'll make this more precise later when we discuss the *Fundamental Theorem of Arithmetic*.

Definition.

A **prime number** is an integer $p > 1$ whose only positive divisors are 1 and p. An integer greater than 1 which is not prime is **composite**.

Remark.

The integer n is composite if and only if there exists an integer a such that $a \mid n$ and $1 < a < n$.

Example.

The integer 7 is prime because its only positive factors are 1 and 7, whereas the integer 9 is composite because it is divisible by 3. ◀

Lemma.

Every integer greater than 1 is divisible by at least one prime.

Proof.

We'll prove the result by induction. To begin with, the result is true for $n = 2,$ since 2 is prime.

Take $n > 2$, and assume the result is true for all integers greater than 1 but less than n. We want to show that the result holds for n. If n is prime, it's divisible by a prime --- namely itself! So suppose n is composite. Then n has a positive factor a other than 1 and n. Suppose $n = ab$.

If $a > n$, then since $b \geq 1,$ We get $n = ab > n.1 = n,$ which is a contradiction. Thus, $a \leq n$, and since $a \neq n$, we have in fact $a < n$. Since $a \neq 1$, we get

$$1 < a < n.$$

By the induction hypothesis, a has a prime factor p.

But $p|a$ and $a|n$ implies $p|n$, so n has a prime factor as well. This shows that the result is true for all $n > 1$ by induction.◀

Theorem. (Euclid)

There are infinitely many prime numbers.

Proof.

Suppose on the contrary that there are only finitely many primes $p_1, p_2, \ldots, p_n$.

Look at $(p_1 \cdot p_2 \cdot \ldots \cdot\ p_n) + 1$.

This number is not divisible by any of the primes $p_1, p_2, \ldots, p_n$, because it leaves a remainder of 1 when divided by any of them. But the previous lemma says that every number greater than 1 is divisible by a prime. This contradiction implies that there can't be finitely many primes --- that is, there are infinitely many. ◀

If you are trying to factor a number *n*, you do not need to try dividing by all the numbers from 1 to *n*: It's enough to go up to $\sqrt{n}$. This is the idea of the next lemma.

Lemma.

Every composite number has a proper factor less than or equal to its square root.

Proof.

Suppose n is composite. We can write $n = ab$, where $1 < a, b < n$. If both $a, b > \sqrt{n}$, then $n = \sqrt{n}\sqrt{n} < a.b = n$.

This contradiction shows that at least one of a, b must be less than or equal to $\sqrt{n}$.◀

From the above theorem, it follows that an integer is prime if it is not divisible by any prime less than or equal to its square root. This leads to the brute-force algorithm known as **trial division**.

To use trial division we divide n by all primes not exceeding $\sqrt{n}$ and conclude that n is prime if it is not divisible by any of these primes.

In fact, you can adapt the preceding proof to show that a composite number must have a *prime* factor less than or equal to its square root.

For an arbitrary number that is several hundred digits in length, it may be impossible with current technology to determine whether the number is prime. In fact, many **cryptographic systems** depend on the difficulty of factoring large numbers.

Example.

Show that 101 is prime.

Solution.

The only primes not exceeding $\sqrt{101}$ are 2, 3, 5, and 7. Because 101 is not divisible by 2, 3, 5, or 7 (the quotient of 101 and each of these integers is not an integer), it follows that 101 is prime. ■

Example.

Show that 127 is prime.

Solution.

To see whether 127 is prime, I only need to see if it has a prime factor$\leq \sqrt{127} \approx 11.27$. You can do the arithmetic to verify that 127 isn't divisible by 2, 3, 5, 7, or 11.

Hence, it must be prime. ■

Because every integer has a prime factorization, it would be useful to have a procedure for finding this prime factorization. Consider the problem of finding the prime factorization of n. Begin by dividing n by successive primes, starting with the smallest prime, 2. If n has a prime factor, then by the above theorem a prime factor p not exceeding $\sqrt{n}$ will be found. So, if no prime factor not exceeding $\sqrt{n}$ is found, then n is prime. Otherwise, if a prime factor p is found, continue by factoring n/p. Note that n/p has no prime factors less than p. Again, if n/p has no prime factor greater than or equal to p and not exceeding its square root, then it is prime. Otherwise, if it has a prime factor q, continue by factoring $n/(pq)$. This procedure is continued until the factorization has been reduced to a prime. This procedure is illustrated in the following example.

Example.

Find the prime factorization of 7007.

Solution.

To find the prime factorization of 7007, first perform divisions of 7007 by successive primes, beginning with 2. None of the primes 2, 3, and 5 divides 7007. However, 7 divides 7007, with $7007/7 = 1001$. Next, divide 1001 by successive primes, beginning with 7. It is immediately seen that 7 also divides 1001, because $1001/7 = 143$. Continue by dividing 143 by successive primes, beginning with 7. Although 7 does not divide 143, 11 does divide 143, and $143/11 = 13$. Because 13 is prime, the procedure is completed. It follows that $7007 = 7 \cdot 1001 = 7 \cdot 7 \cdot 143 = 7 \cdot 7 \cdot 11 \cdot 13$.

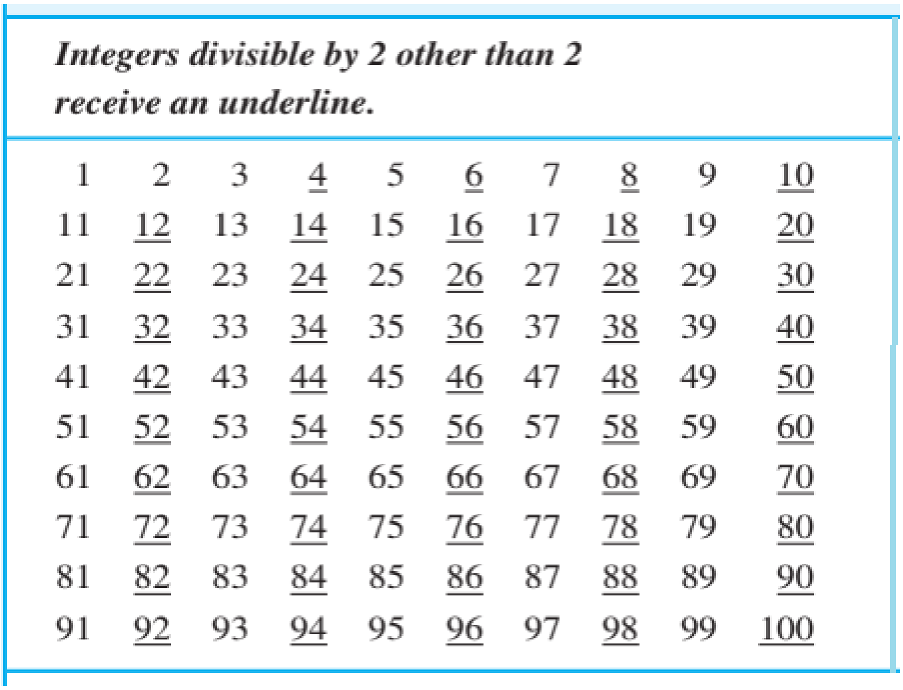
Consequently, the prime factorization of 7007 is $7 \cdot 7 \cdot 11 \cdot 13 = 7^2 \cdot 11 \cdot 13$. ■

Example. (The Sieve of Eratosthenes)

Note that composite integers not exceeding 100 must have a prime factor not exceeding 10. Because the only primes less than 10 are 2, 3, 5, and 7, the primes not exceeding 100 are these four primes and those positive integers greater than 1 and not exceeding 100 that are divisible by none of 2, 3, 5, or 7.

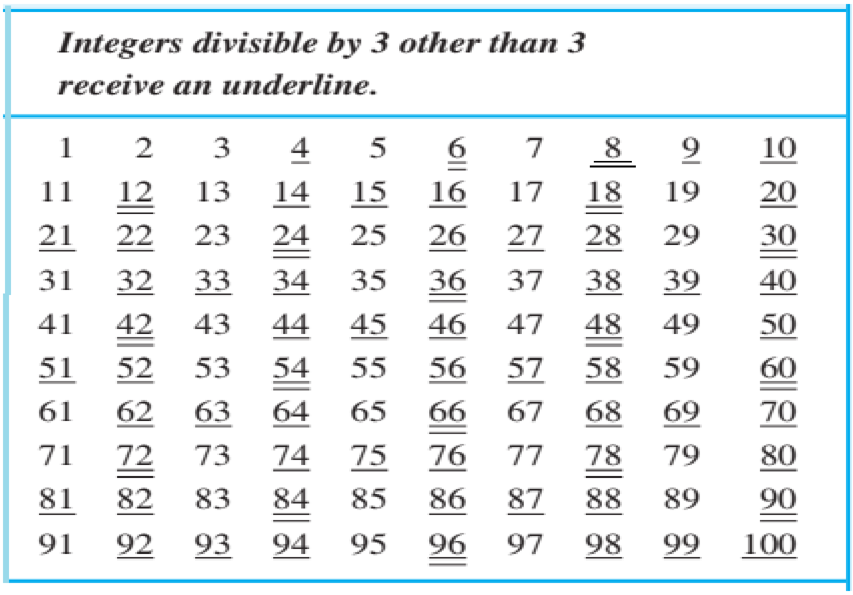
The sieve of Eratosthenes is used to find all primes not exceeding a specified positive integer. For instance, the following procedure is used to find the primes not exceeding 100. We begin with the list of all integers between 1 and 100. To begin the sieving process, the integers that are divisible by 2, other than 2, are deleted. Because 3 is the first integer greater than 2 that is left, all those integers divisible by 3, other than 3, are deleted. Because 5 is the next integer left after 3, those integers divisible by 5, other than 5, are deleted. The next integer left is 7, so those integers divisible by 7, other than 7, are deleted. Because all composite integers not exceeding 100 are divisible by 2, 3, 5, or 7, all remaining integers except 1 are prime. In the table, the panels display those integers deleted at each stage, where each integer divisible by 2, other than 2, is underlined in the first panel, each integer divisible by 3, other than 3, is underlined in the second panel, each integer divisible by 5, other than 5, is underlined in the third panel, and each integer divisible by 7, other than 7, is underlined in the fourth panel. The integers not underlined are the primes not exceeding 100. We conclude that the primes less than

100 are 2, 3, 5, 7, 11, 13, 17, 19, 23, 29, 31, 37, 41, 43, 47, 53, 59, 61, 67, 71, 73, 79, 83, 89, and 97.



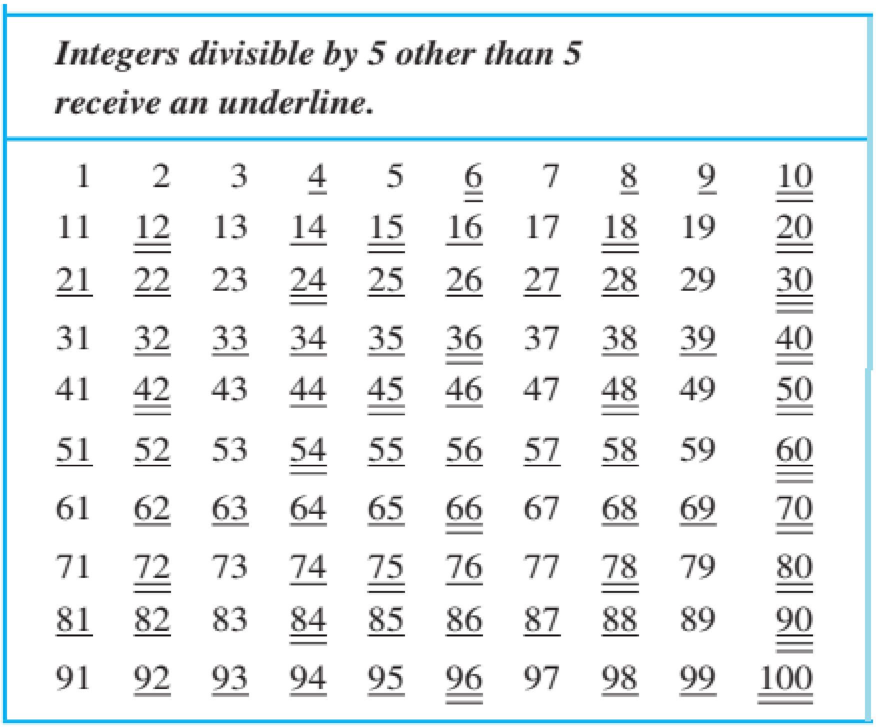
Integers divisible by 2 other than 2 receive an underline.

1	2	3	4	5	6	7	8	9	10
11	12	13	14	15	16	17	18	19	20
21	22	23	24	25	26	27	28	29	30
31	32	33	34	35	36	37	38	39	40
41	42	43	44	45	46	47	48	49	50
51	52	53	54	55	56	57	58	59	60
61	62	63	64	65	66	67	68	69	70
71	72	73	74	75	76	77	78	79	80
81	82	83	84	85	86	87	88	89	90
91	92	93	94	95	96	97	98	99	100



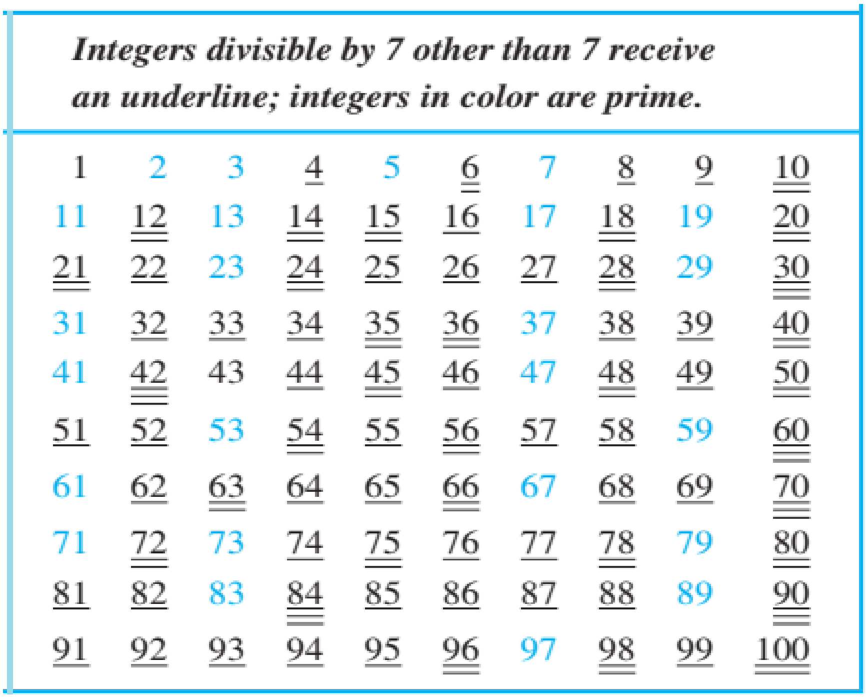
Integers divisible by 3 other than 3 receive an underline.

1	2	3	4	5	6	7	8	9	10
11	12	13	14	15	16	17	18	19	20
21	22	23	24	25	26	27	28	29	30
31	32	33	34	35	36	37	38	39	40
41	42	43	44	45	46	47	48	49	50
51	52	53	54	55	56	57	58	59	60
61	62	63	64	65	66	67	68	69	70
71	72	73	74	75	76	77	78	79	80
81	82	83	84	85	86	87	88	89	90
91	92	93	94	95	96	97	98	99	100



Integers divisible by 5 other than 5 receive an underline.

1	2	3	4	5	6	7	8	9	10
11	12	13	14	15	16	17	18	19	20
21	22	23	24	25	26	27	28	29	30
31	32	33	34	35	36	37	38	39	40
41	42	43	44	45	46	47	48	49	50
51	52	53	54	55	56	57	58	59	60
61	62	63	64	65	66	67	68	69	70
71	72	73	74	75	76	77	78	79	80
81	82	83	84	85	86	87	88	89	90
91	92	93	94	95	96	97	98	99	100



Integers divisible by 7 other than 7 receive an underline; integers in color are prime.

1	2	3	4	5	6	7	8	9	10
11	12	13	14	15	16	17	18	19	20
21	22	23	24	25	26	27	28	29	30
31	32	33	34	35	36	37	38	39	40
41	42	43	44	45	46	47	48	49	50
51	52	53	54	55	56	57	58	59	60
61	62	63	64	65	66	67	68	69	70
71	72	73	74	75	76	77	78	79	80
81	82	83	84	85	86	87	88	89	90
91	92	93	94	95	96	97	98	99	100

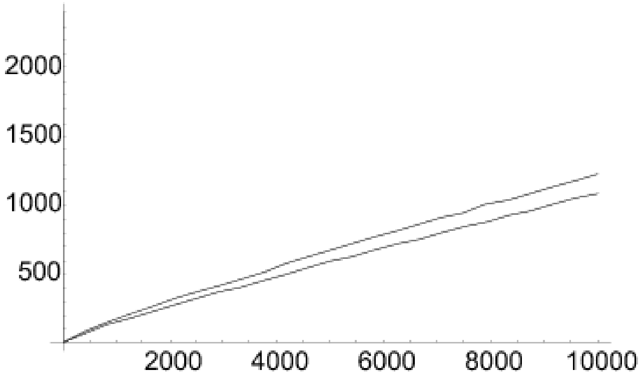
Integers in color are prime.■

I showed above that there are infinitely many primes. How are they distributed? That is, are they evenly distributed, or do they get "sparser" as you look at bigger and bigger integers?

● The Prime Number Theorem gives an asymptotic estimate for $\pi(x)$, the number of primes less than or equal to x. It says:

$$\lim_{x\to\infty} \frac{\pi(x)}{\frac{x}{\ln x}} = 1.$$

The picture below was generated by *Mathematica*, the symbolic mathematics program. It shows the graphs of $\pi(x)$ and $\frac{x}{\ln x}$.



The graph of $\pi(x)$ is on top and the graph of $\frac{x}{\ln x}$ is on the bottom. On the other hand, there are "lots" of composite numbers around. For example,

$$1001!+\ 2,1001!\ +3,1001!\ +4,\ldots,1001!+\ 1001$$

is a run of 1000 consecutive composite numbers. You can use the same method to generate runs of composite numbers of any length.

Example.

Use the Prime Number Theorem to estimate the number of primes less than 1000000. By the Prime Number Theorem, $\pi(1000000) \approx \frac{1000000}{\ln 1000000} \approx 72382.$

The actual number of primes less than 1000000 is $\pi(1000000) = 78498.$■

On the other hand, many problems concerning the distribution of primes are unsolved. For example, there are primes that come in pairs such as 11 and 13, or 71 and 73. These are called **twin primes**.

Question: (Twin Prime Conjecture)

Are there infinitely many twin primes?

There are enormously large twin primes known.

The largest known in 2001 were

$$318032361.2^{107001} \pm 1,$$

They are numbers having 32220 digits! The Twin Prime Conjecture is still unresolved: A proof was announced in 2004, but a gap was found, and the question remains open.

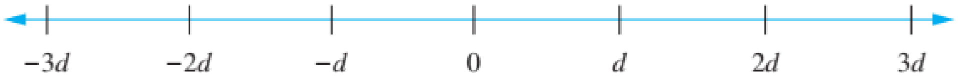
5.3 Divisibility

When one integer is divided by a second nonzero integer, the quotient may or may not be an integer. For example, $12/3 = 4$ is an integer, whereas $11/4 = 2.75$ is not. This leads to the following definition.

Definition.

If a and b are integers with $a = 0$, we say that *a divides b* if there is an integer *c* such that $b = ac$, or equivalently, if is an integer. When *a* divides *b* we say that *a* is a *factor* or *divisor* of *b*, and that *b* is a *multiple* of *a*. The notation $a|b$ denotes that *a* divides *b*. We write $a \nmid b$ when a does not divide b. ■

In the following figure a number line indicates which integers are divisible by the positive integer *d*.



Example.

Determine whether $3|7$ and whether $3|12$.

Solution.

We see that $3|7$, because $7/3$ is not an integer. On the other hand, $3|12$ because $12/3 = 4$.■

Example.

Let *n* and *d* be positive integers. How many positive integers not exceeding *n* are divisible by *d*?

Solution.

The positive integers divisible by d are all the integers of the form dk, where k is a positive integer. Hence, the number of positive integers divisible by d that do not exceed *n* equals the number of integers *k* such that $0 < dk \leq n$, or with $0 < k \leq n/d$. So, there are $\lfloor n/d \rfloor$ positive integers not exceeding *n* that are divisible by d.■

Theorem.

Let a, b, and *c* be integers, where $a \neq 0$. Then

(i) if $a \mid b$ and $a \mid c$, then $a \mid (b + c)$;

(ii) if $a \mid b$, then $a \mid bc$ for all integers *c*;

(iii) if $a \mid b$ and $b \mid c$, then $a \mid c$.

Proof.

We will give a direct proof of (i).

Suppose that $a \mid b$ and $a \mid c$. Then, there are integers s and t with $b = as$ and $c = at$. Hence, $b + c = as + at = a(s + t)$. Therefore, a divides $b + c$. This establishes part (i) of the theorem.■

Corollary.

If $a, b,$ and c are integers, where $a \neq 0$, such that $a|b$ and $a|c$, then $a|mb + nc$ whenever m and n are integers.

Proof.

We will give a direct proof. By part (ii) of the above theorem we see that $a|mb$ and $a|nc$ whenever m and n are integers. By part (i) of the above theorem it follows that $a|mb + nc$.■

Theorem.

Let a, b, c be integers. If $a|b$ and $b|a + c$, then $a|c$.

Proof.

Let a, b, and c be integers. Suppose $a|b$ and $b|a + c$. There exist integers d and e such that $ad = b$ and $be = a + c$.

Substitute: $(ad)e = a + c$.

Rearrange: $a(de - 1) = c$.

Therefore, $a \mid c$.■

Definition.

An integer u is a *unit* if $u|1$.The only units are 1 and –1.

Theorem.

If u and v are units, then $\boldsymbol{uv}$ is a unit.

Proof.

Let u and v be units. There exist integers r and s such that $ur = 1$ and $vs = 1$. Therefore, $(ur)(vs) = 1$.

Rearrange: $(uv)(rs) = 1$. Therefore, uv is a unit. ◀

Theorem.

Let a and b be integers. If $a|b$ and $b|a$, $\frac{a}{b}$ and $\frac{b}{a}$ are units

Proof.

Let a and b be integers. Suppose $a|b$ and $b|a$.There exist integers c and d such that $ac = b$ and $bd = a$.Therefore, $acd = bd = a$. So, $cd = 1$.Thus, c and d are units. ◀

Corollary.

If $a|b$ and $b|a$, then $a = b$ or $a =- b$.

5.4 The Division Algorithm

Theorem.

Let n and d be integers, $d \neq 0$. Then there exist unique integers q and r such that $n = qd + r$ and $0 \leq r < d$. q is the *quotient* and r is the *remainder*.

Example.

What are the quotient and remainder when 101 is divided by 11?

Solution.

We have $101 = 11 \cdot 9 + 2$. Hence, the quotient when 101 is divided by 11 is 9, and the remainder is 2. ■

Example.

What are the quotient and remainder when -11 is divided by 3?

Solution.

We have $-11 = 3(-4) + 1$. Hence, the quotient when -11 is divided by 3, and the remainder is 1. Note that the remainder cannot be negative. Consequently, the remainder is not -2, even though $-11 = 3(-3) - 2$, because $r = -2$ does not satisfy $0 \leq r < 3$. ■

♣Note that the integer a is divisible by the integer d if and only if the remainder is zero when a is divided by d.

Example.

Prove that for any integer n, $n^3 – n$ is a multiple of 6.

Proof.

Divide n by 6 to get q and r: $n = 6q + r$, $0 \le r < 6$.

Substitute: $n^3 – n = (6q + r)^3 – (6q + r)$.

Expand and rearrange:

$$n^3 – n = 6(36q^3 + 18q^2r + 3qr^2 – q) + (r^3 – r).$$

Therefore, $6 \mid (n^3 – n)$ if and only if $6 \mid (r^3 – r)$.

Consider the 6 possible cases:

Case 1: $r = 0$. $r^3 – r = 0^3 – 0 = 0 = 6{\cdot}0$.

Case 2: $r = 1$. $r^3 – r = 1^3 – 1 = 0 = 6{\cdot}0$.

Case 3: $r = 2$. $r^3 – r = 2^3 – 2 = 6 = 6{\cdot}1$.

Case 4: $r = 3$. $r^3 – r = 3^3 – 3 = 24 = 6{\cdot}4$.

Case 5: $r = 4$. $r^3 – r = 4^3 – 4 = 60 = 6{\cdot}10$.

Case 6: $r = 5$. $r^3 – r = 5^3 – 5 = 120 = 6{\cdot}20$.

In every case, $6 \mid (r^3 – r)$.

Therefore, $6 \mid (r^3 – r)$ in general.

Therefore, $6 \mid (n^3 – n)$ for all integers n.■

5.5 Greatest Common Divisors

●The ***greatest common divisor*** $\gcd(m, n)$ of integer m and n is the largest integer which divides both m and n.

●The greatest common divisor can be found using the ***Euclidean algorithm***, which is a process of repeated division.

●The greatest common divisor $\gcd(m, n)$ of m and n is a ***linear combination*** of m and n.

●m and n are ***relatively prime*** if $\gcd(m, n) = 1$.

Definition.

The **greatest common divisor** of two integers (not both zero) is the largest integer which divides both of them.

If a and b are integers (not both 0), the greatest common divisor of a and b is denoted $\gcd(a, b)$.

The greatest common divisor of two integers, not both zero, exists because the set of common divisors of these integers is nonempty and finite. One way to find the greatest common divisor of two integers is to find all the positive common divisors of both integers and then take the largest divisor. This is done in the following examples

Later, a more efficient method of finding greatest common divisors will be given.

Example.

What is the greatest common divisor of 24 and 36?

Solution.

The positive common divisors of 24 and 36 are 1, 2, 3, 4, 6, and 12. Hence, $\gcd(24, 36) = 12$. ■

Example.

$\gcd(4, 6) = 2, \gcd(17, 17) = 17, \gcd(42, 0) = 42,$
$\gcd(12, -15) = 3$. ■

Example.

What is the greatest common divisor of 17 and 22?

Solution.

The integers 17 and 22 have no positive common divisors other than 1, so that $\gcd(17, 22) = 1$. ■

You were probably able to do the last examples by factoring the numbers in your head. For instance, to find $\gcd(4, 6)$, you see that 2 is the only integer bigger than 1 which divides both 4 and 6.

The problem with this approach is that it requires that you factor the numbers. However, once the numbers get too large --- currently, "too large" means "on the order of several hundred digits long" --- this approach to finding the greatest common divisor won't work. Fortunately, the **Euclidean algorithm** computes the greatest common divisor of two numbers without factoring the numbers. I'll discuss it after I state and prove some elementary properties.

Proposition.

Let a and b be integers, not both 0.

(a) $\gcd(a, b) \geq 1$,

(b) $\gcd(a, b) = \gcd(|a|, |b|)$,

(c) $\gcd(a, b) = \gcd(a + kb, b)$ for any integer k.

Proof.

(a) Since $1|a$ and $1|b$, then $\gcd(a, b)$ must be at least as big as 1.

(b) $x|a$ if and only if $x|-a$; that is, a and $-a$ have the same factors. But $|a|$ is either a or $-a$, so a and $|a|$ have the same factors. Likewise, b and $|b|$ have the same factors. Therefore, x is a common factor of a and b if and only if it's a common factor of $|a|$ and $|b|$.

Hence, $\gcd(a,b) = \gcd(|a|,|b|)$.

(c) First, if x is a common factor of a and b, then $x|a$ and $x|b$.

Then $x|kb$, so $x|a+kb$.

Thus we have that x is a common factor of $a+kb$ and b.

Likewise, if x is a common factor of $a+kb$ and b, then $x|a+kb$ and $x|b$.

Hence, $x|(a+kb)-kb=a$.

Thus, x is a common factor of a and b.

Therefore, these two sets are the same:

$$\left\{\begin{matrix}\text{common factors}\\ \text{of } a \text{ and } b\end{matrix}\right\} = \left\{\begin{matrix}\text{common factors}\\ \text{of } a+kb \text{ and } b\end{matrix}\right\}$$

Since the two sets are the same, their largest elements are the same.

The largest element of the first set is $\gcd(a,b)$, while the largest element of the second set is $\gcd(a+kb,b)$.

Therefore, $\gcd(a,b) = \gcd(a+kb,b)$.■

Example.

Part (c) of the proposition says that the greatest common divisor remains unchanged if you add or subtract a multiple of one of the numbers from the other. You can often use this to simplify computations of greatest common divisors. For example,

$$\gcd(998,996) = \gcd(998 - 996,996) = \gcd(2,996).$$

Now $\gcd(2,996)|2$, and the only positive integers which divide 2 are 1 and 2. So $\gcd(2,996)$ is either 1 or 2. But 2 and 996 are obviously both divisible by 2, so $\gcd(2,996) = 2$. Therefore, $\gcd(998,996) = 2$.■

Example.

Prove that if $n \in \mathbb{Z}$, then $\gcd(3n + 4, n + 1) = 1$.

Proof.

By part (c) of the above proposition, we get

$$\gcd(3n + 4, n + 1) = \gcd\big((3n + 4) - 3(n + 1), n + 1\big)$$
$$= \gcd(1, n + 1).$$

Now, $\gcd(1, n + 1)|1$. But the only positive integer which divides 1 is 1. So, $\gcd(1, n + 1) = 1$.

Therefore, $\gcd(3n + 4, n + 1) = 1$.■

♣Because it is often important to specify that two integers have no common positive divisor other than 1, we have the following definition.

Definition.

a and b are **relatively prime** if $\gcd(a, b) = 1$.

Example.

49 and 54 are relatively prime, but 25 and 105 are not.■

Proposition.

If $d = \gcd(m, n)$, then $\gcd\left(\frac{m}{d}, \frac{n}{d}\right) = 1$.

Proof.

Let $m = da$ and $n = db$. Then $\gcd\left(\frac{m}{d}, \frac{n}{d}\right) = \gcd(a, b)$.

Let $p > 0$ and $p|a,\ p|b$. Then we can find e and f such that $a = pe$ and $b = pf$. Thus, $m = dpe$ and $n = dpf$. This shows that dp is a common factor of m and n. Since d is the *greatest* common factor, then $d \geq dp$. Therefore, $1 \geq p$. So, $p = 1$ (since p was a positive integer).

We've proven that 1 is the *only* positive common factor of a and b. Therefore, 1 is the greatest common factor of a and b: $\gcd\left(\frac{m}{d}, \frac{n}{d}\right) = \gcd(a, b) = 1$. ◀

5.6 The Euclidean Algorithm.

Before describing the Euclidean algorithm, we will show how it is used to find $\gcd(91, 287)$.

First, divide 287, the larger of the two integers, by 91, the smaller, to obtain

$$287 \ = \ 91 \ \cdot \ 3 \ + \ 14.$$

Any divisor of 91 and 287 must also be a divisor of

$$287 - 91 \cdot 3 = 14.$$

Also, any divisor of 91 and 14 must also be a divisor of

$$287 \ = \ 91 \ \cdot \ 3 \ + \ 14.$$

Hence, the greatest common divisor of 91 and 287 is the same as the greatest common divisor of 91 and 14.

This means that the problem of finding $\gcd(91, 287)$ has been reduced to the problem of finding $\gcd(91, 14)$.

Next, divide 91 by 14 to obtain

$$91 = 14 \cdot 6 + 7.$$

Because any common divisor of 91 and 14 also divides $91 - 14 \cdot 6 = 7$ and any common divisor of 14 and 7 divides 91, it follows that $\gcd(91, 14) \ = \ \gcd(14, 7)$.

Continue by dividing 14 by 7, to obtain $14 = 7 \cdot 2$.

Because 7 divides 14, it follows that $\gcd(14, 7) \ = \ 7$.

Furthermore, because $\gcd(287, 91) = \gcd(91, 14) = \gcd(14, 7) = 7$, the original problem has been solved.

►We now describe how the Euclidean algorithm works in generality. We will use successive divisions to reduce the problem of finding the greatest common divisor of two positive integers to the same problem with smaller integers, until one of the integers is zero.

♣Begin with a pair of nonnegative integers $\{m, n\}$, not both 0. (The absolute value property we stated earlier shows that there's no harm in assuming the integers are nonnegative.)

1. If one of the numbers is 0, the other is the greatest common divisor of the pair. (Stop.)
2. Otherwise, apply the **Division Algorithm** to write $m = qn + r$,where $0 \leq r < n$.
3. Replace the pair $\{m, n\}$ with the pair $\{n, r\}$.
4. Go to step 1.

At each step, both elements are ≥ 0, and each pass through step 3 decreases the second element. Since the second element always gets smaller, but can't be negative, Well-Ordering implies that algorithm must terminate in an $\{x, 0\}$ pair (in step 2) after a finite number of steps.

I get the next pair of numbers by subtracting a multiple of one of the previous numbers from the other. Therefore, each pair of numbers has the same greatest common divisor as the previous pair. Considering the whole chain of pairs, it follows that the original pair of numbers and the last pair of numbers have the same greatest common divisor.

The original pair of numbers is $\{m, n\}$, and their greatest common divisor is $\gcd(m, n)$.The last pair of numbers is $\{x, 0\}$ and $\gcd(x, 0) = x$. Thus, $\gcd(m, n) = x$--- in words, the greatest common divisor is the last nonzero remainder.

The **Euclidean algorithm** is based on the following result about greatest common divisors and the division algorithm.

Lemma.

Let $a = bq + r$, where a, b, q, and r are integers. Then $\gcd(a, b) = \gcd(b, r)$.

Proof.

If we can show that the common divisors of a and b are the same as the common divisors of b and r, we will have shown that $\gcd(a, b) = \gcd(b, r)$, because both pairs must have the same greatest common divisor. So suppose that d divides both a and b. Then it follows that d also divides $a - bq = r$. Hence, any common divisor of a and b is also a common divisor of b and r.

Likewise, suppose that d divides both b and r. Then d also divides $bq + r = a$. Hence, any common divisor of b and r is also a common divisor of a and b. Consequently, $\gcd(a, b) = \gcd(b, r)$.◀

Example.

Use the Euclidean algorithm to compute:

$$\gcd(124, 348).$$

Solution.

Here what the algorithm above says. You start with the original numbers. Think of them as the first two "remainders". At each step, you divide the next-to-the-last remainder by the last remainder. You stop when you get a remainder of 0.

Here are the divisions:

$348 = 2 \cdot 124 + 100,$
$124 = 1 \cdot 100 + 24,$
$100 = 4 \cdot 24 + 4,$
$24 = 6 \cdot 4 + 0.$

(Start by dividing the bigger number by the smaller number, or else you'll just waste a step.)

It's easier to remember this visually by arranging the computations in a table. Compare the numbers above to the numbers in the following table:

a	q
348	-
124	2
100	1
24	4
4	6

(The next remainder is 0, so I didn't write it.) The successive remainders go in the a-column. The successive quotients go in the q-column. The greatest common divisor is the **last nonzero remainder**, so $\gcd(348, 124) = 4$.

Later on, I'll add another column to this table when I discuss the **Extended Euclidean algorithm**.■

Example.

Use the Euclidean algorithm to compute:

$$\gcd(482, 288).$$

Solution.

a	q
482	
288	1
194	1
94	2
6	15
4	1
2	2

From the table, we see that $\gcd(482, 288) = 2.$■

Example.

Use the Euclidean algorithm to compute:

$$\gcd(414, 662).$$

Solution.

Successive uses of the division algorithm give:

$$662 = 414 \cdot 1 + 248$$

$$414 = 248 \cdot 1 + 166$$
$$248 = 166 \cdot 1 + 82$$
$$166 = 82 \cdot 2 + 2$$
$$82 = 2 \cdot 41.$$

Hence, $\gcd(414, 662) = 2$, because 2 is the last nonzero remainder. ■

Example.

You can also take the greatest common divisor of more than two numbers. For instance, $\gcd(42,105,91) = 7$. To compute the greatest common divisor of more than two divisors, just compute the greatest common divisor two numbers at a time. For example, $\gcd(42, 105) = 21$, so $\gcd(42,105,91) = \gcd\big((42,105), 91\big) = \gcd(21,91) = 7$. ■

5.7 gcds as Linear Combinations

♣The next result is **extremely** important, and is often used in proving things about greatest common divisors. First, I'll recall a definition from linear algebra.

Definition.

If x and y are numbers, **a linear combination** of x and y (with integer coefficients) is a number of the form

$ax + by$, where a and b are integers.

Example.

$29 = 2.10 + 1.9$ shows that 29 is a linear combination of 10 and 9 and $7 = (-2). 10 + 3.9$ shows that 7 is a linear combination of 10 and 9 as well.■

Theorem (c).

$\mathrm{gcd}(\mathrm{m}, \mathrm{n})$ is the smallest positive linear combination of m and n. In particular, there are integers a and b (not necessarily unique) such that

$$\mathrm{gcd}(\mathrm{m}, \mathrm{n}) = \mathrm{am} + \mathrm{bn}.$$

Example.

We showed above that $\mathrm{gcd}(348, 124) \ = \ 4.$

The theorem says that there are integers a and b such that

$4 = a \cdot 348 + b \cdot 124.$

In fact, $4 = 5 \cdot 348 + (-14) \cdot 124.$

This combination is not unique.

For example, $4 = 129 \cdot 348 + (-362) \cdot 124.$■

To Find a and b such that 4 is a linear combination of 384 and 124, we use the backward substations as follows:

$4 = 100 - 4 \cdot 24$

$= 100 - 4 \cdot (124 - 1.100)$

$= -4 \cdot 124 + 5 \cdot 100$

$= -4 \cdot 124 + 5 \cdot (348 - 2 \cdot 124)$

$= 5 \cdot 348 + (-14) \cdot 124.$

We'll give a few easy corollaries before proving the theorem.

Corollary.

If $d|m$ and $d|n$, then $d|\gcd(m, n)$.

Proof.

$\gcd(m, n) = am + bn$ for some integers a and b.

Therefore, if $d|m$ and $d|n$, then $d|(am + bn) = \gcd(m, n)$.◀

♣This says that the greatest common divisor is not only "greatest" in terms of *size*; it's also "greatest" in the sense that any other common factor must *divide* it.

Corollary.

m and n are relatively prime if and only if

$am + bn = 1$ for some integers a and b.

Proof.

Necessity.

Suppose m and n are relatively prime. Then

$\gcd(m, n) \; = \; 1$. By Theorem (c),

$\gcd(m, n) = am + bn$ for some integers a and b.

Therefore, $am + bn = 1$ for some integers a and b.

Sufficiency.

Suppose $am + bn = 1$ for some integers a and b. This says that 1 is a positive linear combination of m and n, so (since 1 is the smallest positive integer) it's the *smallest* positive linear combination of m and n. By Theorem (c), this implies that 1 is the greatest common divisor, and m and n are relatively prime.◀

Proof of Theorem (c).

We'll use the Euclidean algorithm. At each step in the Euclidean algorithm, we replace an old pair of numbers with a new pair of numbers. The proof will go this way.

(a) The first two numbers m and n are linear combinations of m and n.

(b) At each step, if the old numbers are linear combinations of m and n, then so are the new numbers.

(c) By (a) and (b), the last two numbers in the algorithm must be linear combinations of m and n.

(d) The last two numbers in the algorithm are $\gcd(m, n)$ and 0. So, $\gcd(m, n)$ is a linear combination of m and n.

Of these four steps, all are clear except the second.

So here is the proof of step (b).

Suppose that my **old** numbers are $\{x, y\}$, and suppose that they're linear combinations of m and n:

$$x = am + bn \text{ and } y = cm + dn.$$

To do the Euclidean algorithm we divide x by y:

$$x = qy + r, \text{ where } 0 \leq r < y.$$

The **new** numbers are

$\{y, r\} = \{y, x - qy\}$

$$= \{cm + dn, (am + bn) - q(cm + dn)\}$$

$$= \{cm + dn, (a - qc)m + (b - qd)n\}$$

Each of the new numbers is a linear combination of m and n. This proves step (b), and the four steps above show that $\gcd(m, n)$ is a linear combination of m and n.

Next, we have to show that it's the *smallest positive linear combination* of m and n.

Suppose p is a positive linear combination of m and n: $p = am + bn$ for some integers a and b.

$\gcd(m, n)|m$ and $\gcd(m, n)|n$, so $\gcd(m, n)|p$. Both of these numbers are positive, so $\gcd(m, n) \leq p$. Since $\gcd(m, n)$ is smaller than any positive linear combination of m and n, $\gcd(m, n)$ must be the *smallest* positive linear combination of m and n.◀

Example.

$(42,105) = 21$, so the theorem asserts that the set of all linear combinations of 42 and 105 --- that is, the set of all numbers of the form $42a + 105b$... is

$$\ldots, -42, -21, 0, 21, 42, 63, \ldots.$$

Notice that the greatest common divisor is the smallest positive element of this set.■

5.8 The Fundamental Theorem of Arithmetic

●The **Fundamental Theorem of Arithmetic** says that every integer greater than 1 can be factored uniquely into a product of primes.

●**Euclid's lemma** says that if a prime divides a product of two numbers, it must divide at least one of the numbers.

●**The least common multiple** $\mathrm{lcm}[\mathrm{a}, \mathrm{b}]$ of nonzero integers *a* and *b* is the smallest positive integer divisible by both *a* and *b*.

Theorem.(Fundamental Theorem of Arithmetic)

Every integer greater than 1 can be written in the form

$$p_1^{n_1} p_2^{n_2} \dots p_k^{n_k}$$

Where $n_i \geq 0$ and the p_i's are distinct primes. The factorization is unique, except possibly for the order of the factors.

Example.

$$4312 = 2.2156 = 2 \cdot 2 \cdot 1078 = 2 \cdot 2 \cdot 2 \cdot 529$$

$$= 2 \cdot 2 \cdot 2 \cdot 7 \cdot 77 = 2 \cdot 2 \cdot 2 \cdot 7 \cdot 7 \cdot 11.$$

That is, $4312 = 2^3 \cdot 7^2 \cdot 11.$■

We need a couple of lemmas in order to prove the uniqueness part of the Fundamental Theorem. In fact, these lemmas are useful in their own right.

Lemma.

If $m|pq$ and $\gcd(m,p) = 1$, then $m|q$.

Proof.

We write $1 = \gcd(m,p) = \mathrm{am} + \mathrm{bp}$ for some integers a and b. Then $\mathrm{q} = \mathrm{amq} + \mathrm{bpq}$.

Now, $m|amq$ and $m|bpq$ (since $m|pq$).

So $m|(\mathrm{amq} + \mathrm{bpq}) = \mathrm{q}$.◀

Lemma.

If p is prime and $p|a_1a_2 \cdots a_n$, then $p|a_i$ for some i.

For $n = 2$, the result says that if p is prime and $p|ab$, then $p|a$ or $p|b$. This is often called **Euclid's lemma**.

Proof.

Do the case $n = 2$ first. Suppose $p|a_1a_2$ and suppose $p \nmid a_1$. I must show $p|a_2$.Since $\gcd(p,a_1)|p$, and p is prime, we have $\gcd(p,a_1) = 1$ or $\gcd(p,a_1) = p$. If $\gcd(p,a_1) = p$, then $p = \gcd(p,a_1)|a_1$, which contradicts $p \nmid a_1$. Therefore, $\gcd(p, a_1) = 1$. By the above lemma, $p|a_2$. This establishes the result for $n = 2$.

Assume $n > 2$, and assume the result is true when p divides a product of $a_i's$ with less than n factors.

Suppose that $p|a_1a_2 \cdots a_n$. Grouping the terms, I have

$$p|(a_1a_2 \cdots a_{n-1})a_n$$

By the case $n = 2$, either $p|a_1a_2 \cdots a_{n-1}$ or, $p|a_n$. If $p|a_n$, I'm done. Otherwise, if $p|a_1a_2 \cdots a_{n-1}$, then p divides one of $a_1, a_2, \cdots, a_{n-1}$, by induction. In either case, I've shown that p divides one of the $a_i's$, which completes the induction step and the proof.◀

Proof.(Fundamental Theorem of Arithmetic)

First, I'll use induction to show that every integer greater than 1 can be expressed as a product of primes.

$n = 2$ is prime, so the result is true for $n = 2$.

Suppose $n > 2$, and assume every number less than n can be factored into a product of primes. If n is prime, I'm done. Otherwise, n is composite, so I can factor n as $n = ab$, where $1 < a, b < n$. By induction, a and b can be factored into primes. Then $n = ab$ shows that n can, too.

Now I'll prove the uniqueness part of the Fundamental Theorem. Suppose that

$$p_1^{m_1} p_2^{m_2} \dots p_j^{m_j} = q_1^{n_1} q_2^{n_2} \dots q_k^{n_k}$$

Here the p's are distinct primes, the q's are distinct primes, and all the exponents are greater than or equal to 1.

I want to show that$= k$, and that each $p_a^{m_a}$ is $q_b^{n_b}$ for some b --- that is, $p_a = q_b$ and $m_a = n_b$.

Look at p_1. It divides the left side, so it divides the right side. By the Euclid's lemma, $p_1 \middle| q_i^{n_i}$ for some i. But $q_i^{n_i}$ is $q_i \dots q_i$ (n_i times), so again by the Euclid's lemma, $p_1 | q_i$. Since p_1 and q_i are prime, $p_1 = q_i$.

To avoid a mess, renumber the q's so q_i becomes $q_1$and vice versa. Thus, $p_1 = q_i$, and the equation reads

$$p_1^{m_1} p_2^{m_2} \dots p_j^{m_j} = q_1^{n_1} q_2^{n_2} \dots q_k^{n_k}$$

If $m_1 > n_1$, cancel $p_1^{n_1}$ from both sides, leaving

$$p_1^{m_1 - n_1} \dots p_j^{m_j} = q_2^{n_2} \dots q_k^{n_k}.$$

This is impossible, since now p_1 divides the left side, but not the right.

For the same reason $m_1 < n_1$is impossible.

It follows that $m_1 = n_1$. So I can cancel the p_1's off both sides, leaving

$$p_2^{m_2} \dots p_j^{m_j} = q_2^{n_2} \dots q_k^{n_k}.$$

Keep going. At each stage, I pair up a power of a p with a power of a q, and the preceding argument shows the powers are equal. I can't wind up with any primes left over at the end, or else I'd have a product of primes equal to 1. So everything must have paired up, and the original factorizations were the same (except possibly for the order of the factors).◀

Definition.

The ***least common multiple*** of nonzero integers a and b is the smallest positive integer divisible by both a and b. The least common multiple of a and b is denoted $\operatorname{lcm}[a, b]$.

For example, $\operatorname{lcm}[6,4] = 12, \operatorname{lcm}[33,15] = 165$.

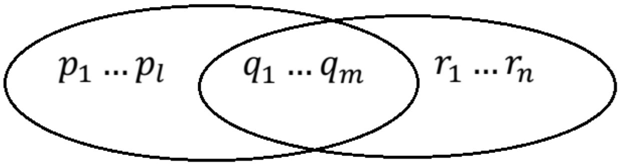
Here's an interesting fact that is easy to derive from the Fundamental Theorem:

$$\operatorname{lcm}[a, b] \cdot \gcd(a, b) = ab.$$

Factor a and b in products of primes, but write out all the powers (e.g. write 2^3 as $2 \cdot 2 \cdot 2$):

$$a = p_1 \dots p_l q_1 \dots q_m, \quad b = q_1 \dots q_m r_1 \dots r_n.$$

Here the q's are the primes a and b have in common, and the p's and r's don't overlap.



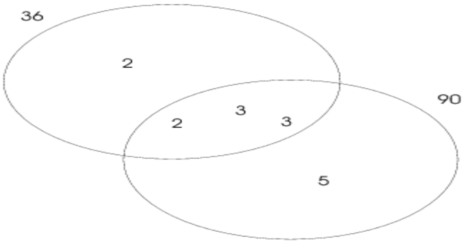
From the picture, $\gcd(a, b) = q_1 \dots q_m$,

$\operatorname{lcm}[a, b] = p_1 \dots p_l q_1 \dots q_m r_1 \dots r_{n,}$

$ab = p_1 \dots p_l q_1 \dots q_m q_1 \dots q_m r_1 \dots r_n$

Thus, $\operatorname{lcm}[a, b]\gcd(a, b) = ab$.

Here's how this result looks for 36 and 90:



$\gcd(36,90) = 18,$

$\operatorname{lcm}[36,90] = 180$ and

$36.90 = 32400 = 18 \cdot 180.$■

5.9 The Chinese Remainder Theorem

●The ***Chinese Remainder Theorem*** gives solutions to systems of congruences with relatively prime moduli.

●The solution to a system of congruences with relatively prime moduli may be produced using a *formula* by computing modular inverses, or using an *iterative procedure* involving successive substitution.

The ***Chinese Remainder Theorem*** says that certain systems of simultaneous congruences *with different moduli* have solutions. The idea embodied in the theorem was apparently known to Chinese mathematicians a long time ago --- hence the name.

I'll begin by collecting some useful lemmas.

Lemma 1.

Let m and $a_1, a_2, \cdots, a_n$ be positive integers. If m is relatively prime to each of $a_1, a_2, \cdots, a_n$, then it is relatively prime to their product $a_1 a_2 \cdots a_n$.

Proof.

If $\gcd(m, a_1 a_2 \cdots a_n) \neq 1$, then there is a prime p which divides both m and $a_1 a_2 \cdots a_n$.

Since $p|a_1a_2\cdots a_n$, we have p must divide a_i for some i *by Euclid's lemma*. Now p divides both m and a_i, so $\gcd(m, a_i) \neq 1$. This contradiction implies that $\gcd(m, a_1a_2\cdots a_n) = 1$.◀

Example.

6 is relatively prime to 25, to 7, and to 11.

$25.7.11 = 1925$, and $\gcd(6, 1925) = 1$:

a	q
1925	-
6	320
5	1
1	5

.■

I showed earlier that the greatest common divisor $\gcd(a, b)$ of a and b is *greatest* in the sense that it is divisible by any common divisor of a and b. The next result is the analogous statement for least common multiples.

Lemma 2.

Let m and $a_1, a_2, \cdots, a_n$ be positive integers. If m is a multiple of each of $a_1, a_2, \cdots, a_n$, then m is a multiple of $\mathrm{lcm}[a_1, a_2, \cdots, a_n]$.

Proof.

By the Division Algorithm, there are unique numbers q and r such that

$$m = q \cdot \mathrm{lcm}[a_1, a_2, \cdots, a_n] + r,$$

Where $0 \leq r < \mathrm{lcm}[a_1, a_2, \cdots, a_n]$.

Now, a_i divides both m and $\mathrm{lcm}[a_1, a_2, \cdots, a_n]$, so a_i divides r. Since this is true for all i, we have r is a common multiple of the a_i smaller than the *least* common multiple $\mathrm{lcm}[a_1, a_2, \cdots, a_n]$. This is only possible if $r = 0$.Then $m = q \cdot \mathrm{lcm}[a_1, a_2, \cdots, a_n]$, i.e. m is a multiple of $\mathrm{lcm}[a_1, a_2, \cdots, a_n]$.◀

Example.

88 is a multiple of 4 and 22. The least common multiple of 4 and 22 is 44, and 88 is also a multiple of 44.■

Lemma 3.

Let $a_1, a_2,\cdots, a_n$ be positive integers. If $a_1, a_2,\cdots, a_n$ are pairwise relatively prime, then

$$\text{lcm}[a_1, a_2,\cdots, a_n] = a_1 a_2 \cdots a_n$$

Proof.

Induction on n. The statement is trivially true for $n = 1$, so I'll start with $n = 2$. The statement for $n = 2$ follows from the equation ,$\text{lcm}[a, b]\text{gcd}(a, b) = ab$.

$$\text{lcm}[a_1, a_2] = \frac{a_1 a_2}{\text{gcd}(a_1, a_2)} = \frac{a_1 a_2}{1} = a_1 a_2.$$

Now assume $n > 2$, and assume the result is true for n. I will prove that it holds for $n + 1$.

Claim:

$$\text{lcm}[\text{lcm}[a_1, a_2,\cdots, a_n], a_{n+1}] = \text{lcm}[a_1, a_2,\cdots, a_n, a_{n+1}]$$

(Some people take this as an iterative *definition* of $\text{lcm}[a_1, a_2,\cdots, a_n, a_{n+1}]$).

$\text{lcm}[a_1, a_2,\cdots, a_n, a_{n+1}]$ is a multiple of each of $a_1, a_2,\cdots, a_n$, so by Lemma 2 it's a multiple of $\text{lcm}[a_1, a_2,\cdots, a_n]$. It's also a multiple of a_{n+1}, so

$$\text{lcm}[\text{lcm}[a_1, a_2,\cdots, a_n], a_{n+1}] | \text{lcm}[a_1, a_2,\cdots, a_n, a_{n+1}].$$

On the other hand, for $i = 1, \ldots, n$,

$a_i|\operatorname{lcm}[a_1,\cdots,a_n]$

and $\operatorname{lcm}[a_1,\cdots,a_n]|\operatorname{lcm}[\operatorname{lcm}[a_1,\cdots,a_n],a_{n+1}]$.

Therefore, $a_i|\operatorname{lcm}[\operatorname{lcm}[a_1,a_2,\cdots,a_n],a_{n+1}]$.

Obviously,

$$a_{n+1}|\operatorname{lcm}[\operatorname{lcm}[a_1,a_2,\cdots,a_n],a_{n+1}].$$

Thus, $\operatorname{lcm}[\operatorname{lcm}[a_1,a_2,\cdots,a_n],a_{n+1}]$ is a common multiple of all the a_i's. Since $\operatorname{lcm}[a_1,a_2,\cdots,a_n,a_{n+1}]$ is the least common multiple, Lemma 2 implies that

$\operatorname{lcm}[a_1,a_2,\cdots,a_n,a_{n+1}]|\operatorname{lcm}[\operatorname{lcm}[a_1,a_2,\cdots,a_n],a_{n+1}]$.

Since I have two *positive* numbers which divide one another, they're equal:

$$\operatorname{lcm}[\operatorname{lcm}[a_1,a_2,\cdots,a_n],a_{n+1}]=\operatorname{lcm}[a_1,a_2,\cdots,a_n,a_{n+1}]$$

This proves the claim.

Returning to the proof of the induction step, I have

$$\begin{aligned}\operatorname{lcm}[a_1,a_2,\cdots,a_n,a_{n+1}] &= \operatorname{lcm}[\operatorname{lcm}[a_1,a_2,\cdots,a_n],a_{n+1}]\\ &= \operatorname{lcm}[a_1a_2\cdots a_n,a_{n+1}]\\ &= a_1a_2\cdots a_n.\end{aligned}$$

The second equality follows by the induction hypothesis (the statement for n). The third equality follows from Lemma 1 and the result for $n=2$.◀

Example.

6, 25, and 7 are relatively prime (in pairs). The least common multiple is

$$\mathrm{lcm}[6,25,7] = 1050 = 6.25.7.\blacksquare$$

Lemma 4.

Let m be a positive integer and let a, b, and c be integers. If $ac \equiv bc (\mathrm{mod}\, m)$ and $\gcd(c, m) = 1$, then $a \equiv b\ (\mathrm{mod}\, m)$.

Proof.

Because $ac \equiv bc (\mathrm{mod}\, m)$, $m|ac - bc = c(a - b)$. Because $\gcd(c, m) = 1$, it follows that $m|a - b$. We conclude that $a \equiv b(\mathrm{mod}\, m)$.◀

The Chinese Remainder Theorem

Theorem. (The Chinese Remainder Theorem)

Suppose $m_1, m_2, \ldots, m_n$ are pairwise relatively prime (that is, $\gcd(m_i, m_j) = 1$ for $i \neq j$). Then the system of congruences

$$x \equiv a_1 (\operatorname{mod} m_1)$$
$$x \equiv a_2 (\operatorname{mod} m_2)$$
$$\vdots$$
$$x \equiv a_n (\operatorname{mod} m_n)$$

has a unique solution mod $m_1 m_2 \ldots m_n$.

Notation. $x_1 x_2 \ldots, \widehat{x_i}, \ldots, x_n$

means the product $x_1 x_2 \ldots, x_i, \ldots, x_n$with x_i omitted. For example, $x_1 x_2 \ldots, \widehat{x_4}, \ldots, x_6$ means $x_1 x_2 x_3 x_5 x_6$.

This is a convenient (and standard) notation for omitting a single variable term in a product of things.

Proof.

Define $p_k = m_1 \ldots \widehat{m_k} \ldots m_n$.

That is, p_k is the product of the m's with m_k omitted. By Lemma 1, $\gcd(p_k, m_k) = 1$. Hence, there are integer numbers s_k, t_k such that $s_k p_k + t_k m_k = 1$.

In terms of congruences, $s_k p_k \equiv 1(\text{mod}\, m_k)$.

Now let $x = a_1 p_1 s_1 + a_2 p_2 s_2 + \cdots + a_n p_n s_n$.

If $j \neq k$, then $m_k \big| p_j$, so mod m_k all the terms but the k^{th} term die: $x \equiv a_k p_k s_k = a_k .1 \equiv a_k$ $(\text{mod}\, m_k)$

This proves that x is a solution to the system of congruences (and incidentally, gives a formula for x).

Now suppose that x and y are two solutions to the system of congruences.

$$\begin{array}{ccc} x \equiv a_1(\text{mod}\, m_1) & \text{and} & y \equiv a_1(\text{mod}\, m_1) \\ x \equiv a_2(\text{mod}\, m_2) & \text{and} & y \equiv a_2(\text{mod}\, m_2) \\ \vdots & \vdots & \vdots \\ x \equiv a_n(\text{mod}\, m_n) & \text{and} & y \equiv a_n(\text{mod}\, m_n) \end{array}$$

Then $x \equiv a_k \equiv y$ $(\text{mod}\, m_k)$.

So $x - y \; \equiv 0(\text{mod}\, m_k)$ or $m_k | x - y$.

Thus, $x - y$ is a multiple of all the m's, so

$$\text{lcm}[m_1, m_2, \dots, m_n] | x - y.$$

But the m's are pairwise relatively prime. By Lemma 3,

$$m_1 m_2 \dots m_n | x - y, \text{ i.e. } x \equiv y(\text{mod}\, m_1 m_2 \dots m_n).$$

That is, the solution to the congruences is unique mod $m_1 m_2 \dots m_n$.■

Example.

Solve

$$x \equiv 2(\bmod 4);$$

$$x \equiv 7(\bmod 9).$$

Solution.

Since $\gcd(4,9) = 1$, so there is a unique solution mod 36. Following the construction of *x* in the proof,

$$p_1 = 9, \qquad 9 \cdot 1 \equiv 1(\bmod 4), \text{so take } s_1 = 1$$

$$p_2 = 4, \qquad 4 \cdot 7 \equiv 1(\bmod 9), \text{so take } s_2 = 7$$

$$x = a_1 p_1 s_1 + a_2 p_2 s_2 = 18 + 196 = 214$$

$$\equiv 34(\bmod 36).$$ ■

Example.

Solve

$$x \equiv 3(\bmod 4);$$

$$x \equiv 1(\bmod 5);$$

$$x \equiv 2(\bmod 3).$$

Solution.

The moduli are pairwise relatively prime, so there is a unique solution mod 60. This time, I'll solve the system using an iterative method.

$x \equiv 3 (\text{mod } 4)$, so $x = 3 + 4s$

But $x \equiv 1 (\text{mod } 5)$, so $3 + 4s \equiv 1 \; (\text{mod } 5)$,

$4s \equiv 3 \; (\text{mod } 5)$, implies $4.4s \equiv 4.3 \; (\text{mod } 5)$.

Since $4.3 \equiv 32 \; (\text{mod } 5)$, we have $4.4s \equiv 32 \; (\text{mod } 5)$

$s \equiv 2 \; (\text{mod } 5), s = 2 + 5t$.

Hence, $x = 3 + 4s = 3 + 4(2 + 5t) = 11 + 20t$.

Finally, $x \equiv 2 \; (\text{mod } 3)$, so

$$11 + 20t \equiv 2 \; (\text{mod } 3),$$

$20t \equiv -9 \equiv 0 \; (\text{mod } 3), 20t \equiv 0 \; (\text{mod } 3)$,

$20t \equiv 20.0 \; (\text{mod } 3), t \equiv 0 \; (\text{mod } 3)$,

Hence, $t = 3u$.

Now put everything back:

$x = 11 + 20t = 11 + 20(3u) = 11 + 60u$,or

$x \equiv 11 (\text{mod } 60)$.■

Example.

Ahmed keeps balls in his bag. If he divides them into 5 equal groups, 4 are left over. If he divides them into 8 equal groups, 6 are left over. If he divides them into 9 equal groups, 8 are left over. What is the smallest number of balls that he could have?

Solution.

Let x be the number of balls. Then

$$x \equiv 4(\mathrm{mod}\ 5);$$

$$x \equiv 6(\mathrm{mod}\ 8);$$

$$x \equiv 8(\mathrm{mod}\ 9).$$

From $x \equiv 4(\mathrm{mod}\ 5)$, I get$x = 4 + 5a$. Plugging this into the second congruence, I get

$$4 + 5a \equiv 6(\mathrm{mod}\ 8)$$

$$5a \equiv 2(\mathrm{mod}\ 8)$$

$$5 \cdot 5a \equiv 5 \cdot 2(\mathrm{mod}\ 8)$$

$$25a \equiv 10(\mathrm{mod}\ 8),$$

But,

$$10 \equiv 50(\mathrm{mod}\ 8)$$

Then,

$$25a \equiv 50(\mathrm{mod}\ 8)$$

Or,

$$a \equiv 2(\mathrm{mod}\ 8)$$

Hence,$a = 2 + 8b$. Plugging this into$x = 4 + 5a$ gives $x = 4 + 5(2 + 8b) = 14 + 40b$.

Plugging this into the third congruence, I get

$$14 + 40b \equiv 8(\mathrm{mod}\ 9)$$

$$40b \equiv -6 (\text{mod } 9)$$

But,

$$-6 \equiv 120 (\text{mod } 9)$$

So,

$$40b \equiv 120 (\text{mod } 9)$$

Or,

$$b \equiv 3 (\text{mod } 9)$$

Hence, $b = 3 + 9c$. Plugging this into $x = 14 + 40b$ gives $x = 14 + 40(3 + 9c) = 134 + 360c$.

The smallest positive value of x is obtained by setting $c = 0$, which gives $x = 134$. ■

You can sometimes solve a system even if the moduli aren't relatively prime; the criteria are similar to those for solving system of linear Diophantine equations. I'll state the result, but omit the proof.

Theorem.

Consider the system

$$x \equiv a_1 (\text{mod } m_1)$$

$$x \equiv a_2 (\text{mod } m_2)$$

(a) If $\gcd(m_1, m_2) \nmid a_1 - a_2$, there are no solutions.

(b) If $\gcd(m_1, m_2) | a_1 - a_2$, there is a unique solution mod $\operatorname{lcm}[m_1, m_2]$.■

Note that if $\gcd(m_1, m_2) = 1$, case (b) automatically holds, and $\operatorname{lcm}[m_1, m_2] = m_1 m_2$ --- i.e. I get the Chinese Remainder Theorem for $n = 2$.

Example.

Solve $x \equiv 5 (\mathrm{mod}\ 12)$; $x \equiv 11 (\mathrm{mod}\ 18)$.

Solution.

Since $\gcd(12, 18) = 6 | 11 - 5$, there is a unique solution mod $\operatorname{lcm}[12, 18] = 36$. I'll use the iterative method to find the solution. $x \equiv 5 (\mathrm{mod}\ 12)$, so $x = 5 + 12s$.

Since $x \equiv 11 (\mathrm{mod}\ 18)$,

$$5 + 12s \equiv 11 (\mathrm{mod}\ 18), 12s \equiv 6 (\mathrm{mod}\ 18)$$

Now I use my rule for "dividing" congruencies: 6 divides both 12 and 6, and $\gcd(6,18) = 6$, so I can divide through by 6: $2s \equiv 1 (\mathrm{mod}\ 3)$

Multiply by 2, and convert the congruence to an equation:

$$s \equiv 2 (\mathrm{mod}\ 3), s = 2 + 3t.$$

Plug back in:

$x = 5 + 12(2 + 3t) = 29 + 36t, x \equiv 29 (\mathrm{mod}\ 36)$.■

Exercise Set (5)

(1) Let $a, b, c \in \mathbb{Z}$. Prove that:

(a) If $a|b$ and $a|c$, then $a|bx + cy$ for all $x, y \in \mathbb{Z}$.

(b) If $a|b$, then $a|bc$.

(c) If $a|b$ and $b|c$, then $a|c$.

(d) If $a > 0$, $b > 0$ and $a|b$ then $a \leq b$.

(e) If $a|b$, then $|a|\big||b|$.

(f) If $a|b$ and $b|a$ then $a = \pm b$.

(2) Let *a* be positive integer and *b* integer. Prove that there exist unique integers r, q such that:

$$b = qa + r \text{ , where } 0 \leq r < a.$$

(3) Prove that:

If $b = qa + r$, then $\gcd(a, b) = \gcd(a, r)$.

(4) Evaluate $\gcd(6755, 1587645)$ and find *x* and *y* such that $\gcd(6755, 1587645) = 6755x + 1587645y$.

(5) Evaluate $\gcd(123456789, 987654321)$ and find *x* and *y* such that

$$\gcd(123456789, 987654321) = 123456789x + 987654321y$$

(6) Evaluate $\gcd(189, 283, 512)$ and find x, y and z such that $\gcd(189, 283, 512) = 189\,x + 283\,y + 512\,z$.

(7) Find $\gcd(360, -2250)$, $\operatorname{lcm}[360, -2250]$, $\gcd(3799, 7337)$, $\operatorname{lcm}[3799, 7337]$, $\operatorname{lcm}[6,10,14]$, $\operatorname{lcm}[7, 11, 13]$.

(8) Prove that $\sqrt{2}$, $\sqrt[3]{10}$, $\log_{10} 2$ are irrationals.

(9) Let $n < p \leq 2n$. Prove that:

$$n^{\pi(2n)-\pi(n)} \leq 2^{2n} \text{and } \frac{\pi(2n)-\pi(n)}{n} < \frac{2}{\log_2 n}.$$

(10) Let $a, b, c \in \mathbb{Z}$and $n \in \mathbb{Z}^+$. Prove that:

(a) $a \equiv a \pmod{n}$;

(b) If $a \equiv b \pmod{n}$, then $b \equiv a \pmod{n}$;

(c) If $a \equiv b \pmod{n}$, and $b \equiv c \pmod{n}$, then $a \equiv c \pmod{n}$;

(d) If $a \equiv b \pmod{n}$, and $c \equiv d \pmod{n}$, then $a + c \equiv b + d \pmod{n}$, $a - c \equiv b - d \pmod{n}$, and $ac \equiv bd \pmod{n}$.

(e) If $a \equiv b \pmod{n}$ and $m|n$, then $a \equiv b \pmod{m}$.

(11) Let $a, b, c \in \mathbb{Z}$ and $n \in \mathbb{Z}^+$. Prove that:

(a) If $ac \equiv bc \pmod{n}$and $(c, n) = 1$, then $a \equiv b \pmod{n}$.

(b) If $ac \equiv bc \pmod{p}$, then $a \equiv b \pmod{p}$, where p is prime and does not divide c.

(12) Find the least positive integer x such that:

3 divides x with remainder 1; 4 divides x with remainder 2; 5 divides x with remainder 3; (hint: Use the Chinese remainder theorem).

(13) Solve:

a) $19x \equiv 1 \pmod{140}$;

b) $13x \equiv 71 \pmod{380}$.

c) $108x \equiv 171 \pmod{529}$.

(14) Solve the following systems:

a) $x \equiv 2 \pmod{3}, x \equiv 3 \pmod{5}, x \equiv 2 \pmod{7}$.

b) $x \equiv 1 \pmod{3}, x \equiv 2 \pmod{4}, x \equiv 3 \pmod{5}$.

c) $x \equiv 5 \pmod{7}, x \equiv 12 \pmod{15}$,

$$x \equiv 18 \pmod{22}.$$

CHAPTER (VI)

COUNTING

Chapter (VI)

Counting

The Beginning of Mathematics

6.1 The Basics of Counting

Suppose that a password on a computer system consists of six, seven, or eight characters. Each of these characters must be a digit or a letter of the alphabet. Each password must contain at least one digit. How many such passwords are there? The techniques needed to answer this question and a wide variety of other counting problems will be introduced in this section. Counting problems arise throughout mathematics and computer science. For example, we must count the successful outcomes of experiments and all the possible outcomes of these experiments to determine probabilities of discrete events. We need to count the number of operations used by an algorithm to study its time complexity. We will introduce the basic techniques of counting in this section. These methods serve as the foundation for almost all counting techniques.

6.2 Basic Counting Principles

1-The Addition rule or Sum rule

(Principle of disjunction counting)

If a task can be done either in one of n_1 ways or in one of n_2 ways, where none of the set of n_1 ways is the same as any of the set of n_2 ways, then there are $n_1 + n_2$ ways to do the task.

More general:

Let S be a set and $|S|$ denote the number of elements in S. If S is a union of disjoint non-empty subsets

$$A_1, A_2, \dots, A_n,$$

then

$$|S| = |A_1| + |A_2| + \cdots + |A_n|.$$

In the above statement the subsets A_i of S are all disjoint i.e., they have no element in common. If A_i and A_j are two subsets of S, then $A_i \cap A_j = \phi$ for $i \neq j$ and we have $S = A_1 \cup A_2 \cup \dots \cup A_n$ that is each element of S is exactly in one of the subsets A_i. In other words, the subsets $A_1, A_2, \dots, A_n$ is a partition of S.

Example

In a class of 30 students, there are 16 boys and 14 girls $(16 + 14 = 30)$. Of these, 23 persons wear pants and only 7 wear skirts $(23 + 7 = 30)$. On the last exam 20 students received a passing grade, while 10 failed

$(20 + 10 = 30)$. ■

Example

An electronic book of 472 pages has been stored in separate files - one file per page - in two folders. One folder contained 305 files, the other 167 files

$(305 + 167 = 472.)$ ■

Example.

There are 40 students in an algebra class and 40 students in a geometry class. How many different students are in both classes combined?

This problem is not well formulated and cannot be answered unless we are told how many students are taking both algebra and geometry. If there is not student taking both algebra and geometry, then by the sum rule the answer is $40 + 40$. But let us assume that there are 10 students taking both algebra and geometry. Then there are 30 students *only* in algebra, 30 students *only* in geometry, and 10 students in *both* algebra and geometry. Therefore, by the sum rule the total number of students is $30 + 30 + 10 = 70.$ ■

Example.

How many ways can we get a sum of 7 or 11 when two distinguishable dice are rolled?

Solution.

The two dice are distinguishable, therefore the ordered pairs (a, b) and (b, a) are distinct when $a \neq b$, i.e., $(a, b) \neq (b, a)$ for $a \neq b$.

The ordered pairs in which the sum is 7 are:

$(1,6), (2,5), (3,4), (4,3), (5,2), (6,1)$.

These ordered pairs are distinct.

∴ There are 6 ways to obtain the sum 7.

Similarly, the ordered pairs: $(5,6), (6,5)$ are all distinct.

∴ The number of ways in which we get a sum 11 with the two dice is 2.

∴ We can get a sum 7 or 11 with two distinguishable dice in $6 + 2 = 8$ ways.■

Example.

How many ways can we draw a club or a diamond from a pack of cards?

Solution.

There are 13 clubs and 13 diamonds in a pack of cards. The number of ways a club or a diamond may be drawn $13 + 13 = 26$. ■

Example.

In how ways can be drawn an ace or a king from an ordinary deck of playing cards?

Solution.

Number of Aces in a pack $= \mathbf{4}$.

Number of kings in a pack $= \mathbf{4}$.

Number ways an Ace or a king can be drawn from the pack $= \mathbf{4} + \mathbf{4} = \mathbf{8}$. ■

Example.

Suppose that either a member of the mathematics faculty or a student who is a mathematics major is chosen as a representative to a university committee. How many different choices are there for this representative if there are 37 members of the mathematics faculty and 83 mathematics majors and no one is both a faculty member and a student?

Solution.

There are 37 ways to choose a member of the mathematics faculty and there are 83 ways to choose a student who is a mathematics major. Choosing a member of the mathematics faculty is never the same as choosing a student who is a mathematics major because no one is both a faculty member and a student. By the sum rule it follows that there are $37 + 83 = 120$ possible ways to pick this representative. ■

Example.

A student can choose a computer project from one of three lists. The three lists contain 23, 15, and 19 possible projects, respectively. No project is on more than one list. How many possible projects are there to choose from?

Solution.

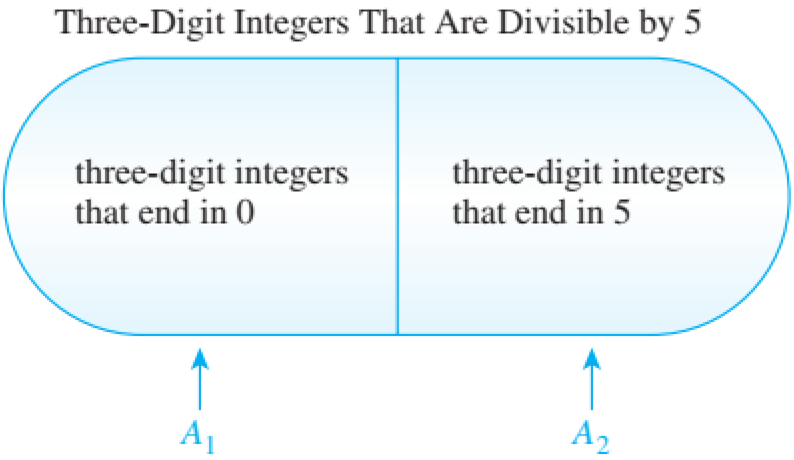
The student can choose a project by selecting a project from the first list, the second list, or the third list. Because no project is on more than one list, by the sum rule there are $23 + 15 + 19 = 57$ ways to choose a project. ■

Example.

How many three-digit integers (integers from 100 to 999 inclusive) are divisible by 5?

Solution.

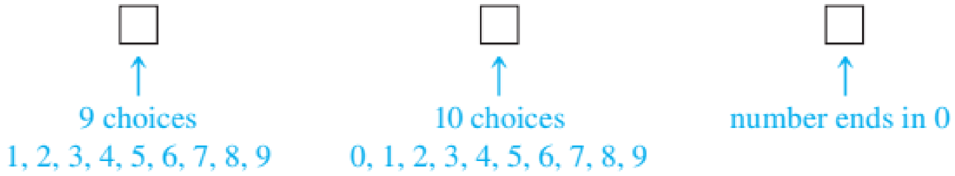
We use the addition rule. Integers that are divisible by 5 end either in 5 or in 0. Thus the set of all three-digit integers that are divisible by 5 can be split into two mutually disjoint subsets A_1 and A_2 as shown in the following figure.


Three-Digit Integers That Are Divisible by 5
three-digit integers that end in 0
three-digit integers that end in 5
A_1
A_2

$A_1 \cup A_2 =$the set of all three-digit integers that are divisible by 5. $A_1 \cap A_2 = \phi$.

Now there are as many three-digit integers that end in 0 as there are possible choices for the left-most and middle digits (because the right-most digit must be a 0). As illustrated below, there are nine choices for the left-most digit (the digits 1 through 9) and ten choices for the middle digit (the digits 0 through 9). Hence

$$|A_1| = 9 \cdot 10 = 90.$$



Similar reasoning shows that there are as many three-digit integers that end in 5 as there are possible choices for the left-most and middle digits, which are the same as for the integers that end in 0. Hence,

$$|A_2| = 9 \cdot 10 = 90.$$

The number of three-digit integers that are divisible by 5 $= |A_1| + |A_2| = 90 + 90 = 180.$ ■

2-Product Rule (The Multiplication Rule)

The Principle of Sequential Counting

Consider the following example. Suppose a computer installation has four input/output units (*A*, *B*, *C*, and *D*) and three central processing units (*X*, *Y*, and *Z*). Any input/output unit can be paired with any central processing unit. How many ways are there to pair an input/output unit with a central processing unit?

To answer this question, imagine the pairing of the two types of units as a two-step

operation:

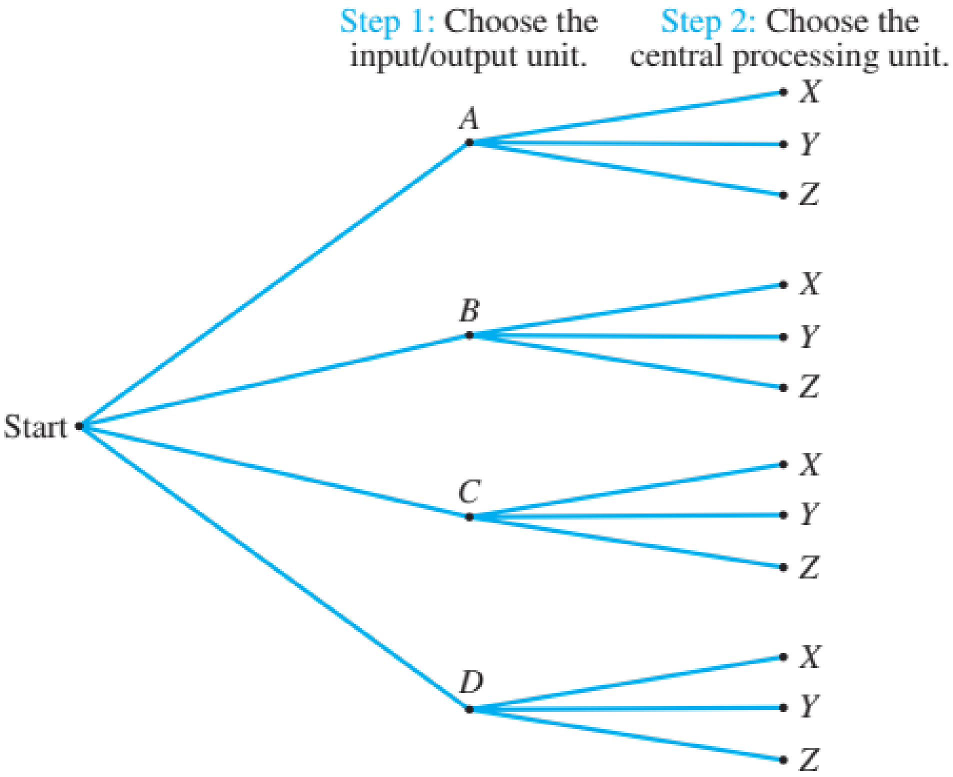
Step 1: Choose the input/output unit.

Step 2: Choose the central processing unit.

The possible outcomes of this operation are illustrated in the possibility tree of the following figure.

The topmost path from “root” to “leaf” indicates that input/output unit *A* is to be paired with central processing unit *X*. The next lower branch indicates that input/output unit *A* is to be paired with central processing unit *Y*. And so forth.

Thus the total number of ways to pair the two types of units is the same as the number of branches of the tree, which is $\mathbf{3+3+3+3=4\cdot 3=12}$.


Step 1: Choose the input/output unit.
Step 2: Choose the central processing unit.
Start
A
B
C
D
X
Y
Z
X
Y
Z
X
Y
Z
X
Y
Z

The idea behind this example can be used to prove the following rule.

THE PRODUCT RULE Suppose that a procedure can be broken down into a sequence of two tasks. If there are $\boldsymbol{n_1}$ ways to do the first task and for each of these ways of doing the first task, there are $\boldsymbol{n_2}$ ways to do the second task, then there are $\boldsymbol{n_1} \times \boldsymbol{n_2}$ ways to do the procedure.

In general:

If an operation consists of k steps and the first step can be performed in $\boldsymbol{n_1}$ ways,

the second step can be performed in $\boldsymbol{n_2}$ ways [regardless of how the first step was performed],

$$\vdots$$

the kth step can be performed in $\boldsymbol{n_k}$ ways [regardless of how the preceding steps were performed], then the entire operation can be performed in $\boldsymbol{n_1} \times \boldsymbol{n_2} \times \ldots \times \boldsymbol{n_k}$ ways.

◀

Example.

There are two drawers. One contains 12 shirts, the other 7 neckties. There are $84 = 12 \times 7$ ways to combine a shirt and a necktie. It is possible to examine the drawers sequentially: first-second, first-second... It is also possible to form combinations using two hands: left for a shirt, right for a necktie. As long as all possible combinations shirt/necktie have been counted, the exact procedure is of no consequence.■

Example.

A test consists of 6 multiple-choice questions. Each question has 4 possible answers. There are $4 \times 4 \times 4 \times 4 \times 4 \times 4 = 4^6$ ways to answer all 6 questions. ■

Example.

There are boxes in a postal office labelled with an English letter (out of 26 English characters) and a positive integer not exceeding 80. How many boxes with different labels are possible?

Solution.

The procedure of labelling boxes consists of two successive stages. In the first stage we assign 26 different English letters, and in the second stage we assign 80 natural numbers (the second stage does *not* depend on the outcome of the first stage). Thus by the multiplication rule we have $26 \times 80 = 2080$ different labels. ■

Example.

How many different bit strings are there of length five?

Solution.

We have here a procedure that assigns two values (i.e., zero or one) in five stages. Therefore, by the multiplication rule we have $2^5 = 32$ different strings. ■

Example.

How many possible outcomes are there when we roll a pair of dice, one red and one green?

Solution.

The red die can land in any one of six ways and for each of their six ways, the green die can also land in six ways. The number of possible outcomes when two dice are rolled $= \mathbf{6} \times \mathbf{6} = \mathbf{36}$.■

Example.

In how many different ways one can answer all the questions of a true-false test consisting of 4 questions?

Solution.

There are two ways of answering each of the 4 questions. So by product rule the number of ways in which all the 4 questions can be answered$= \mathbf{2} \times \mathbf{2} \times \mathbf{2} \times \mathbf{2} = \mathbf{16}$. ■

Example.

Find the number $\boldsymbol{n}$ of license plates that can be made where each plate contains two distinct letters followed by three different digits.

Solution.

First letter can be printed in 26 different ways. Since the second letter must be different from the first, we have 25 contains for the second letter. Similarly the first digit can be printed in 10 ways, the second digit in the license plate can be printed in 9 ways and the third in 8 ways. So, the number of license plates that can be printed, so that each plate contains two distinct letters follower by three different digits $26 \times 25 \times 10 \times 9 \times 8 = 4{,}68{,}000.$ ■

Example.

How many functions are there from a set with m elements to a set with n elements?

Solution.

A function corresponds to a choice of one of the $\boldsymbol{n}$ elements in the codomain for each of the $\boldsymbol{m}$ elements in the domain. Hence, by the product rule there are

$\boldsymbol{n} \cdot \boldsymbol{n} \cdot \cdots \cdot \boldsymbol{n} = \boldsymbol{n}^{\boldsymbol{m}}$ functions from a set with $\boldsymbol{m}$ elements to one with $\boldsymbol{n}$ elements. For example, there are $\mathbf{5^3 = 125}$ different functions from a set with three elements to a set with five elements. ■

Example.

A certain personal identification number (PIN) is required to be a sequence of any four symbols chosen from the 26 uppercase letters in the Roman alphabet and the 10 digits.

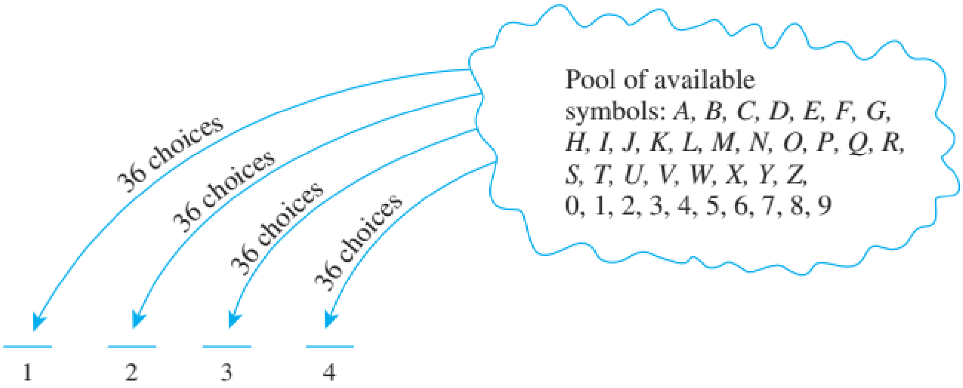
a. How many different PINs are possible if repetition of symbols is allowed?

b. How many different PINs are possible if repetition of symbols is not allowed?

c. What is the probability that a PIN does not have a repeated symbol assuming that all PINs are equally likely?

Solution.

a. Some possible PINs are RCAE, 3387, B92B, and so forth. You can think of forming a PIN as a 4-step operation where each step involves placing a symbol into one of 4 positions, as shown below.


Pool of available symbols: *A, B, C, D, E, F, G, H, I, J, K, L, M, N, O, P, Q, R, S, T, U, V, W, X, Y, Z,* 0, 1, 2, 3, 4, 5, 6, 7, 8, 9
36 choices
36 choices
36 choices
36 choices
1
2
3
4

Step 1: Choose a symbol to place in position 1.

Step 2: Choose a symbol to place in position 2.

Step 3: Choose a symbol to place in position 3.

Step 4: Choose a symbol to place in position 4.

There is a fixed number of ways to perform each step, namely 36, regardless of how preceding steps were performed. And so, by the multiplication rule, there are $\mathbf{36 \cdot 36 \cdot 36 \cdot 36 = 36^4 = \ 1,679,616}$ PINs in all.

b. Again think of forming a PIN as a four-step operation: Choose the first symbol, then the second, then the third, and then the fourth. There are 36 ways to choose the first symbol, 35 ways to choose the second (since the first symbol cannot be used again), 34 ways to choose the third (since the first two symbols cannot be reused), and 33 ways to choose the fourth (since the first three symbols cannot be reused). Thus, the multiplication rule can be applied to conclude that there are

$$\mathbf{36 \cdot 35 \cdot 34 \cdot 33 = 1,413,720}$$

different PINs with no repeated symbol.

c. By part (b) there are $\mathbf{1,413,720}$ PINs with no repeated symbol, and by part (a) there are $\mathbf{1,679,616}$ PINs in all. So the probability that a PIN chosen at random contains no repeated symbol is $\mathbf{1,679,616/ 1,413,720 \cong 0.8417}$. In other words, approximately 84% of PINs have no repeated symbol. ■

Let us now consider some more sophisticated counting problems in which one must use a mixture of the sum and multiplication rules.

Example.

A valid file name must be six to eight characters long and each name must have at least one digit. How many file names can there be?

Solution.

If N is the total number of valid file names and N_6, N_7 and N_8 are, respectively, file names of length six, seven, and eight, then by the sum rule $N = N_6 + N_7 + N_8$:

Let us first estimate N_6. We compute it in an indirect way using the multiplication rule together with the sum rule. We first estimate the number of file names of length six without the constraint that there must be at least one digit. By the multiplication rule there are $(26 + 10)^6 = 36^6$ file names. Now the number of file names that consists of *only* letters (no digits) is 26^6. We must subtract these since they are not allowed. Therefore (by the sum rule)

$N_6 = 36^6 – 26^6 = 1867866560$:

In a similar way, we compute $N_7= 36^7 – 26^7$; $N_8 = 36^8 – 26^8$. Finally $N = N_6 + N_7 + N_8 = 2684483063360$.■

3-The Subtraction Rule (Inclusion–Exclusion for Two Sets)

Suppose that a task can be done in one of two ways, but some of the ways to do it are common to both ways. In this situation, we cannot use the sum rule to count the number of ways to do the task. If we add the number of ways to do the tasks in these two ways, we get an overcount of the total number of ways to do it, because the ways to do the task that are common to the two ways are counted twice. To correctly count the number of ways to do the two tasks, we must subtract the number of ways that are counted twice. This leads us to an important counting rule.

THE SUBTRACTION RULE If a task can be done in either n_1 ways or n_2 ways, then the number of ways to do the task is $n_1 + n_2$ minus the number of ways to do the task that are common to the two different ways.◀

The subtraction rule is also known as the **principle of inclusion–exclusion**, especially when it is used to count the number of elements in the union of two sets.

Suppose that A and B are sets. Then, there are $|A|$ ways to select an element from A and $|B|$ ways to select an element from B. The number of ways to select an element from A or from B, that is, the number of ways to select an element from their union, is the sum of the number of ways to select an element from A and the number of ways to select an element from B, minus the number of ways to select an element that is in both A and B. Because there are $|A \cup B|$ ways to select an element in either A or in B, and $|A \cap B|$ ways to select an element common to both sets, we have

$$|A \cup B| = |A| + |B| - |A \cap B|. \quad (*)$$

Here's an argument that may appear more rigorous.

- Since $A \cap B$ and $B - A$ are disjoint as are A and $B - A$, moreover

$A \cup B = A \cup (B - A)$

and

$B = (A \cap B) \cup (B - A)$,

it follows from (*) that

$$|A \cup B| = |A| + |B - A|,$$

$$|A \cap B| + |B - A| = |B|,$$

which, when added, yield (*).

- If $A \cap B = \phi$, then $|A \cup B| = |A| + |B|$.

Example.

A computer company receives 350 applications from computer graduates for a job planning a line of new Web servers. Suppose that 220 of these applicants majored in computer science, 147 majored in business, and 51 majored both in computer science and in business. How many of these applicants majored neither in computer science nor in business?

Solution.

To find the number of these applicants who majored neither in computer science nor in business, we can subtract the number of students who majored either in computer science or in business (or both) from the total number of applicants. Let A be the set of students who majored in computer science and B the set of students who majored in business. Then $A \cup B$ is the set of students who majored in computer science or business (or both), and $A \cap B$ is the set of students who majored both in computer science and in business. By the subtraction rule the number of students who majored either in computer science or in business (or both) equals

$$|A \cup B| = |A| + |B| - |A \cap B|$$
$$= 220 + 147 - 51 = 316.$$

We conclude that $350 - 316 = 34$ of the applicants majored neither in computer science nor in business. ■

● The story of course does not end here. What about if there are three sets: A, B, C? For three sets, the **Inclusion-Exclusion Principle** reads

$$
\begin{aligned}
&|\mathrm{A} \cup \mathrm{B} \cup \mathrm{C}| \\
&\qquad\qquad = |\mathrm{A}| + |\mathrm{B}| + |\mathrm{C}| - |\mathrm{A} \cap \mathrm{B}| - |\mathrm{B} \cap \mathrm{C}| \\
&\qquad\qquad - |\mathrm{A} \cap \mathrm{C}| + |\mathrm{A} \cap \mathrm{B} \cap \mathrm{C}|
\end{aligned}
$$

● In the more general case where there are *n* different sets $\mathrm{A_i}$, the formula for the **Inclusion-Exclusion Principle** becomes:

$$
\begin{aligned}
\left|\bigcup_{i=1}^{n} A_i\right| = \sum_{i=1}^{n} |A_i| - \sum_{1\le i<j\le n}^{\square} \left|A_i \cap A_j\right| \\
+ \textstyle\sum_{1\le i<j<k\le n}^{\square} \left|A_i \cap A_j \cap A_k\right| \\
- \cdots + (-1)^{n-1} |\textstyle\bigcap_{i=1}^{n} A_i| \text{ .... (**)}
\end{aligned}
$$

● What does (**) say? On the left is number of elements in the union of *n* sets. On the right, we first count elements in each of the sets separately and add the up, as we already know, if the sets $\mathrm{A_i}$ are not disjoint, some elements will have be counted more than once. Those are the elements that belong to *at least* two of the sets $\mathrm{A_i}$, or the intersections $\mathrm{A_i} \cap \mathrm{A_j}$. We wish to consider every such intersection, but each only once. Since $\mathrm{A_i} \cap \mathrm{A_j} = \mathrm{A_j} \cap \mathrm{A_i}$, to avoid duplications we arbitrarily decide to consider only pairs ($\mathrm{A_i} \cap \mathrm{A_j}$) with $\mathrm{i} < \mathrm{j}$.

● When we subtract the sum of the number of elements in such pairwise intersections, some elements may have been subtracted more than once. Those are the elements that belong to at least three of the sets $\mathrm{A_i}$. We add the sum of the elements of intersections of the sets taken three at a time. (The condition $\mathrm{i < j < k}$ assures that every intersection is counted only once.)

● The process goes on with sums being alternately added or subtracted until we come to the last term - the intersection of all sets $\mathrm{A_i}$. Whether it's added or subtracted depends on *n*: for $\mathrm{n = 2}$ it was subtracted, for $\mathrm{n = 3}$ added - take a clue from here.

● Sets A_i are often taken to be subsets of a larger set X such that each A_i is a collection of elements of X that share some property P_i.

$$\bigcup_{i=1}^{n} A_i$$

is then the subset of X that consists of all elements of X having at least one of the properties P_i. Its complement

$$X - \bigcup_{i=1}^{n} A_i$$

is the set of elements that have none of those properties:

$$X - \bigcup_{i=1}^{n} A_i = \bigcap_{i=1}^{n} (X - A_i^{⬚}) = \bigcap_{i=1}^{n} A_i^c$$

from which

$$\left|\bigcap_{i=1}^{n} A_i^c\right| = |X| - \left|\bigcup_{i=1}^{n} A_i\right|$$

This leads to an additional form of (**)

$$\begin{aligned}\left|\bigcap_{i=1}^{n} A_i^c\right| = |X| - \sum_{i=1}^{n}|A_i| + \sum_{1\le i<j\le n}^{\square}\left|A_i \cap A_j\right| \\ - \sum_{1\le i<j<k\le n}^{\square}\left|A_i \cap A_j \cap A_k\right| \\ + \cdots + (-1)^n\left|\bigcap_{i=1}^{n} A_i\right| \text{ .... (***)}\end{aligned}$$

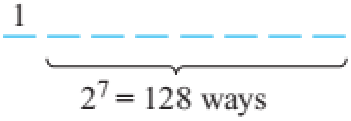
The left-hand side in (***) gives the number of elements of X that have none of the properties P_i.

Example.

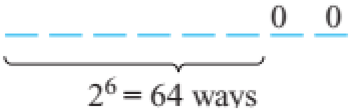
How many bit strings of length eight either start with a 1 bit or end with the two bits 00?

Solution.

We can construct a bit string of length eight that either starts with a 1 bit or ends with the two bits 00, by constructing a bit string of length eight beginning with a 1 bit or by constructing a bit string of length eight that ends with the two bits 00. We can construct a bit string of length eight that begins with a 1 in $2^7 = 128$ ways. This follows by the product rule, because the first bit can be chosen in only one way and each of the other seven bits can be chosen in two ways.

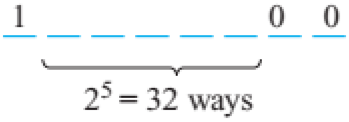


Similarly, we can construct a bit string of length eight ending with the two bits 00, in $2^6 = 64$ ways. This follows by the product rule, because each of the first six bits can be chosen in two ways and the last two bits can be chosen in only one way.



Some of the ways to construct a bit string of length eight starting with a 1 are the same as the ways to construct a bit string of length eight that ends with the two bits 00. There are $2^5 = 32$ ways to construct such a string. This follows by the product rule, because the first

bit can be chosen in only one way, each of the second through the sixth bits can be chosen in two ways, and the last two bits can be chosen in one way.



Consequently, the number of bit strings of length eight that begin with a 1 or end with a 00, which equals the number of ways to construct a bit string of length eight that begins with a 1 or that ends with 00, equals $128 + 64 - 32 = 160$. ■

Example.

A professor in a discrete mathematics class passes out a form asking students to check all the mathematics and computer science courses they have recently taken. He found that, out of a total of 50 students in the class,

30 took precalculus;

16 took both precalculus and Python;

18 took calculus;

8 took both calculus and Python;

26 took Python;

47 took at least one of the three courses.

9 took both precalculus and calculus;

Note that when we write "30 students took precalculus," we mean that the total number of students who took precalculus is 30, and we allow for the possibility that some of these students may have taken one or both of the other courses. If we want to say that 30 students took precalculus only (and not either of the other courses), we will say so explicitly.

a. How many students did not take any of the three courses?

b. How many students took all three courses?

c. How many students took precalculus and calculus but not Python? How many students took precalculus but neither calculus nor Python?

Solution

a. By the difference rule, the number of students who did not take any of the three courses equals the number in the class minus the number who took at least one course. Thus the number of students who did not take any of the three courses is $50 - 47 = 3$.

b. Let

$P =$ the set of students who took precalculus.

$C =$ the set of students who took calculus.

$Y =$ the set of students who took Python.

Then, by the inclusion/exclusion rule,

$$\begin{aligned}|P \cup C \cup Y| &\\ &= |P| + |C| + |Y| - |P \cap C| - |P \cap Y| \\ &- |C \cap Y| + |P \cap C \cap Y|\end{aligned}$$

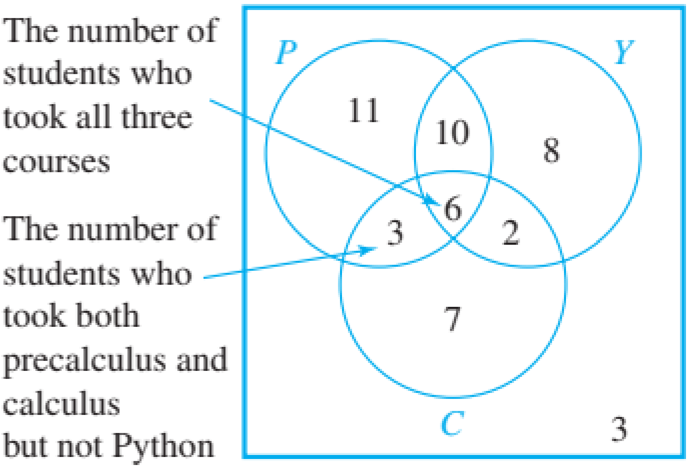
Substituting known values, we get

$47 = 30 + 18 + 26 - 9 - 16 - 8 + |P \cap C \cap Y|$.

Solving for $|P \cap C \cap Y|$ gives $|P \cap C \cap Y| = 6$.

Hence there are six students who took all three courses. In general, if you know any seven of the eight terms in the inclusion/exclusion formula for three sets, you can solve for the eighth term.

c. To answer the questions of part (c), look at the diagram in the following figure.



The number of students who took all three courses

The number of students who took both precalculus and calculus but not Python

Since $|P \cap C \cap Y| = 6$, put the number 6 inside the innermost region. Then work outward to find the numbers of students represented by the other regions of the diagram.

For example, since nine students took both precalculus and calculus and six took all three courses, $9 - 6 = 3$ students took precalculus and calculus but not Python. Similarly, since 16 students took precalculus and Python and six took all three courses, $16 - 6 = 10$ students took precalculus and Python but not calculus. Now the total number of students who took precalculus is 30. Of these 30, three also took calculus but not Python, ten took Python but not calculus, and six took both calculus and Python. That leaves 11 students who took precalculus but neither of the other two courses. A similar analysis can be used to fill in the numbers for the other regions of the diagram. ■

4-Tree diagrams

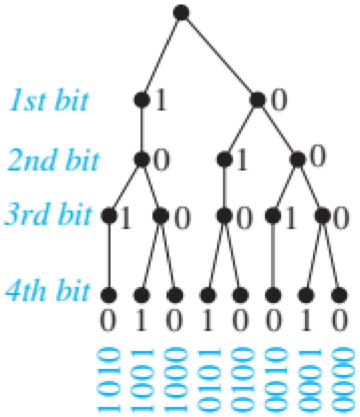
Tree is a structure that consists of a root, branches and leaves. Can be useful to represent a counting problem and record the choices we made for alternatives. The count appears on the leaf nodes. We will study trees later.

Example.

What is the number of bit strings of length 4 that do not have two consecutive ones.

Solution.

The tree diagram in the given figure displays all bit strings of length four without two consecutive 1s. We see that there are eight bit strings of length four without two consecutive 1s. ■

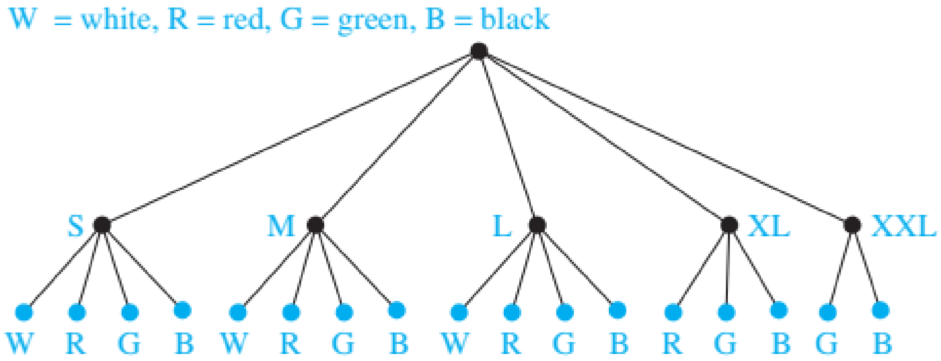


Example.

Suppose that “I Love El-Ahly” T-shirts come in five different sizes: S, M, L, XL, and XXL. Further suppose that each size comes in four colors, white, red, green, and black, except for XL, which comes only in red, green, and black, and XXL, which comes only in green and black. How many different shirts does a souvenir shop have to stock to have at least one of each available size and color of the T-shirt?

Solution.

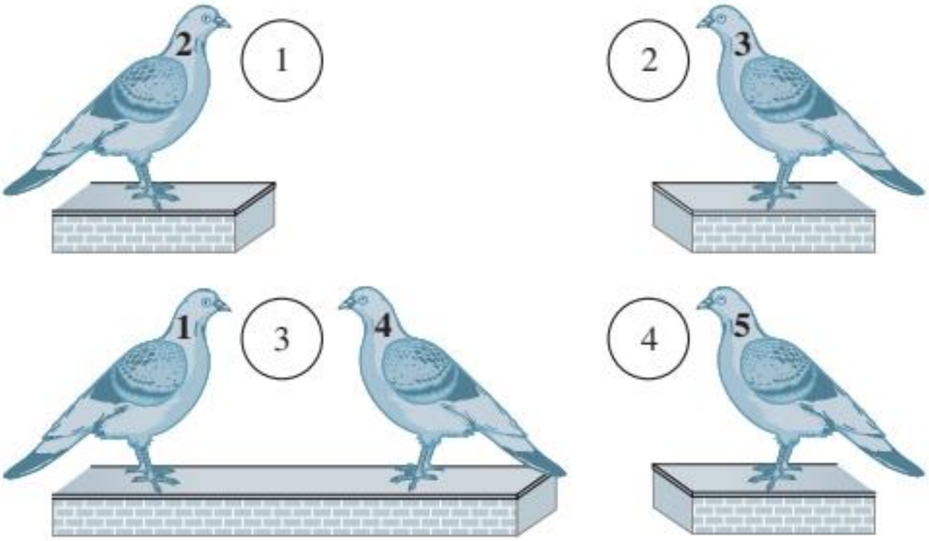
The tree diagram in the following figure displays all possible size and color pairs.


W = white, R = red, G = green, B = black
S
M
L
XL
XXL
W R G B W R G B W R G B R G B G B

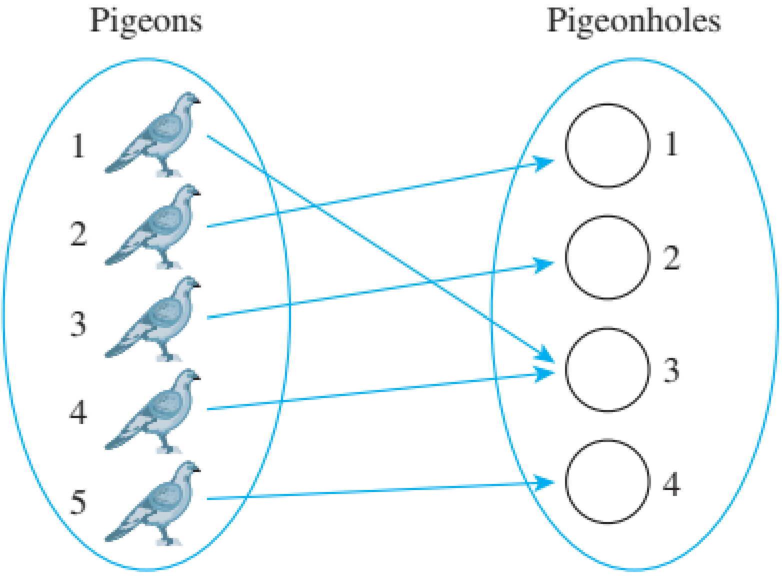
It follows that the souvenir shop owner needs to stock 17 different T-shirts. ■

5-Pigeonhole Principle

The pigeonhole principle states that if n pigeons fly into m pigeonholes and $n > m$, then at least one hole must contain two or more pigeons. This principle is illustrated in the following figures. Illustration (a) shows the pigeons perched next to their holes, and (b) shows the correspondence from pigeons to pigeonholes.



(a)



(b)

Illustration (b) suggests the following mathematical way to phrase the principle.

Pigeonhole Principle

If k is a positive integer and $k + 1$ or more objects are placed into k boxes, then there is at least one box containing two or more of the objects.

Corollary

A function from one finite set to a smaller finite set cannot be one-to-one: There must be at least two elements in the domain that have the same image in the co-domain.◀

The following examples show how the pigeonhole principle is used.

Example.

Among any group of 367 people, there must be at least two with the same birthday, because
there are only 366 possible birthdays. ■

Example.

In any group of 27 English words, there must be at least two that begin with the same letter, because there are 26 letters in the English alphabet. ■

Example.

How many students must be in a class to guarantee that at least two students receive the same score on the final exam, if the exam is graded on a scale from 0 to 100 points?

Solution.

There are 101 possible scores on the final. The pigeonhole principle shows that among any 102 students there must be at least 2 students with the same score. ■

Example.

If 10 pigeons have to fit into 9 pigeonholes, then some pigeonhole gets more than one pigeon. ■

● More generally, if the number of pigeons is greater than the number of pigeonholes, then some pigeonhole gets more than one pigeon.

Example.

Consider a chess board with two of the diagonally opposite corners removed. Is it possible to cover the board with pieces of domino whose size is exactly two board squares?

Solution.

No, it's not possible. Two diagonally opposite squares on a chess board are of the same color. Therefore, when these are removed, the number of squares of one color exceeds by 2 the number of squares of another color. However, every piece of domino covers exactly two squares and these are of different colors.

Every placement of domino pieces establishes a one-to-one correspondence between the set of white squares and the set of black squares. If the two sets have different number of elements, then, by the Pigeonhole Principle, no 1-1 correspondence between the two sets is possible. ■

6-generalized Pigeonhole Principle

For any function f from a finite set X with n elements to a finite set Y with m elements and for any positive integer k, if $km < n$, then there is some $y \in Y$ such that y is the image of at least $k + 1$ distinct elements of X.

Theorem. **THE GENERALIZED PIGEONHOLE PRINCIPLE**

If N objects are placed into k bins then there is at least one bin containing at least $\lceil N/k \rceil$ objects.

Example.

Assume 100 people. Can you tell something about the number of people born in the same month?

Solution

- Yes. There exists a month in which at least $\lceil 100/12 \rceil = \lceil 8.3 \rceil = 9$ people were born. ■

Example.

What is the minimum number of students required in a discrete mathematics class to be sure that at least six will receive the same grade, if there are five possible grades, A, B, C, D, and F?

Solution.

The minimum number of students needed to ensure that at least six students receive the same grade is the smallest integer N such that $\lceil N/5 \rceil = 6$. The smallest such integer is $N = 5 \cdot 5 + 1 = 26$. If you have only 25 students, it is possible for there to be five who have received each grade so that no six students have received the same grade. Thus, 26 is the minimum number of students needed to ensure that at least six students will receive the same grade. ■

Example.

a) How many cards must be selected from a standard deck of 52 cards to guarantee that at least three cards of the same suit are chosen?

b) How many must be selected to guarantee that at least three hearts are selected?

Solution.

a) Suppose there are four boxes, one for each suit, and as cards are selected they are placed in the box reserved for cards of that suit. Using the generalized pigeonhole principle,

we see that if N cards are selected, there is at least one box containing at least $N/4$ cards. Consequently, we know that at least three cards of one suit are selected if $\lceil N/4 \rceil \geq 3$. The smallest integer N such that $\lceil N/4 \rceil \geq 3$ is $N = 2 \cdot 4 + 1 = 9$, so nine cards suffice. Note that if eight cards are selected, it is possible to have two cards of each suit, so more than eight cards are needed. Consequently, nine cards must be selected to guarantee that at least three cards of one suit are chosen. One good way to think about this is to note that after the eighth card is chosen, there is no way to avoid having a third card of some suit.

b) We do not use the generalized pigeonhole principle to answer this question, because we want to make sure that there are three hearts, not just three cards of one suit. Note that in the worst case, we can select all the clubs, diamonds, and spades, 39 cards in all, before we select a single heart. The next three cards will be all hearts, so we may need to select 42 cards to get three hearts. ■

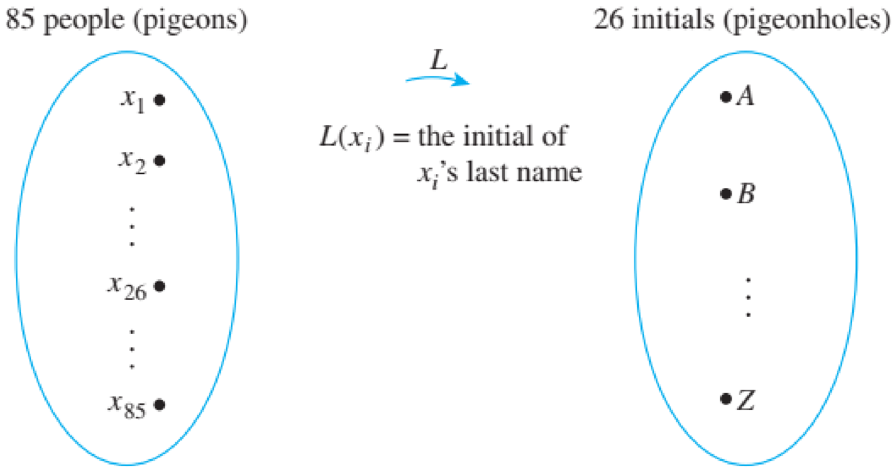
Example.

Show how the generalized pigeonhole principle implies that in a group of 85 people, at least 4 must have the same last initial.

Solution.

In this example the pigeons are the 85 people and the pigeonholes are the 26 possible last initials of their names.

Consider the function L from people to initials defined by the following arrow diagram.


85 people (pigeons)
x_1
x_2
x_{26}
x_{85}
L
$L(x_i)$ = the initial of x_i's last name
26 initials (pigeonholes)
A
B
Z

Since $3 \cdot 26 = 78 < 85$, the generalized pigeonhole principle states that some initial must be the image of at least four $(3 + 1)$ people. Thus at least four people have the same last initial. ■

7-Permutations and Combinations

In computer science one often needs to know in how many ways one can arrange certain objects (e.g., how many inputs are there consisting of ten digits?). To answer these questions, we study here permutations and combinations – the simplest arrangements of objects.

Definition.

A **permutation** of a set of distinct objects is *an ordered arrangements* of these objects. An ordered arrangements of r elements of a set is called an r**-permutation**.

Example.

Let $S = \{1, 2, 3\}$. The ordered arrangement $(3, 1, 2)$ is a permutation of S. The ordered arrangement $(3, 2)$ is a 2-permutation of S.■

Example.

In how many ways can we select three students from a group of five students to stand in line for a picture? In how many ways can we arrange all five of these students in a line for a picture?

Solution

First, note that the order in which we select the students matters. There are five ways to select the first student to stand at the start of the line. Once this student has been selected, there are four ways to select the second student in the line. After the first and second students have been selected, there are three ways to select the third student in the line. By the product rule, there are $5 \cdot 4 \cdot 3 = 60$ ways to select three students from a group of five students to stand in line for a picture.

To arrange all five students in a line for a picture, we select the first student in five ways, the second in four ways, the third in three ways, the fourth in two ways, and the fifth in one way.

Consequently, there are $5 \cdot 4 \cdot 3 \cdot 2 \cdot 1 = 120$ ways to arrange all five students in a line for

a picture. ■

In how many ways may one count a set of *n* elements? Or, which is the same, how many permutations are there of (a set of) *n* elements?

Definition.

The number of permutations of a set of n *elements is denoted and defined by* n! *(pronounced* n *factorial.)*

The number of r-permutations of a set with n *elements is denoted and defined by*

$$P(n; r) = \frac{n!}{(n-r)!}$$

Theorem.

For all integer $n > 0,\ n! = n \cdot (n-1)!$.

Thus n! is the number of ways to count a set of *n* elements. As we saw, 2! = 2. Obviously, 1! = 1, 3! = 6. Indeed, there are just six ways to count three elements:

1. $(1, 2, 3)$
2. $(1, 3, 2)$
3. $(2, 1, 3)$
4. $(2, 3, 1)$
5. $(3, 1, 2)$
6. $(3, 2, 1)$

How many ways are there to count an empty set, the set with 0 elements? (Note that {0} contains one element thus is not empty. The empty set contains no elements at all - {}.) Since there is nothing to count the question is *In how many ways can one do nothing?* A mathematical answer to this is *just one:* $0! = 1$.

An aside

There is just one way to do nothing so that $0! = 1$. However, the result of this activity is nothing or, in math parlance, 0. You may enjoy the following question.

Guess the next number in the following sequence

$$0, 1, 2, 720!$$

Answer to the problem

$720! = (6!)! = ((3!)!)!$,

i.e. three followed by three factorials.

$2 = 2! = (2!)!$,

i.e. two followed by two factorials.

$1 = 1!$

and finally, $0! =$ 0 followed by zero factorials - a result of doing nothing.

The answer then is 4!!!! The number is quite big (how big?). So that computing it would take a lot of effort.

Here is another way to do this. Look at the six permutations of a 3-element set. Let's try mimicking this for a set of n elements. There are n ways to select the first element. For each of these, by definition, the remaining $(n-1)$ elements can be counted in $(n-1)!$ ways. Therefore, there are $n \cdot (n - 1)!$ ways to count an n-element set.

Example.

Let $S = \{1, 2, 3, 4, 5\}$. The ordered arrangement $(4, 2, 1, 5, 3)$ is a permutation of S.

$(3, 1, 4)$ is a 3-permutation of S.■

Example.

Let $S = \{a, b, c\}$. The 2- permutations of S are the ordered arrangements a, b; a, c; b, a; b, c; c, a; and c, b. Consequently, there are six 2-permutations of this set with three elements. There are always six 2-permutations of a set with three elements. There are three ways to choose the first element of the arrangement. There are two ways to choose the second element of the arrangement, because it must be different from the first element. Hence, by the product rule, we see that $P\ (3, 2)\ =\ 3\ \cdot\ 2\ =\ 6$. the first element. By the product rule, it follows that $P\ (3, 2)\ =\ 3\ \cdot\ 2\ =\ 6$.■

Example.

How many ways are there to select a first-prize winner, a second-prize winner, and a third-prize winner from 100 different people who have entered a contest?

Solution.

Because it matters which person wins which prize, the number of ways to pick the three prize winners is the number of ordered selections of three elements from a set of 100 elements, that is, the number of 3-permutations of a set of 100 elements. Consequently, the answer is $P\,(100,3) = 100 \cdot 99 \cdot 98 = 970{,}200$. ■

Example.

Suppose that there are eight runners in a race. The winner receives a gold medal, the second place finisher receives a silver medal, and the third-place finisher receives a bronze medal. How many different ways are there to award these medals, if all possible outcomes of the race can occur and there are no ties?

Solution.

The number of different ways to award the medals is the number of 3-permutations of a set with eight elements. Hence, there are $P\,(8,3) = 8 \cdot 7 \cdot 6 = 336$ possible ways to award the medals. ■

Example.

Suppose that a saleswoman has to visit eight different cities. She must begin her trip in a specified city, but she can visit the other seven cities in any order she wishes. How many possible orders can the saleswoman use when visiting these cities?

Solution.

The number of possible paths between the cities is the number of permutations of seven elements, because the first city is determined, but the remaining seven can be ordered arbitrarily. Consequently, there are $7! = 5040$ ways for the sales woman to choose her tour. If, for instance, the saleswoman wishes to find the path between the cities with minimum distance, and she computes the total distance for each possible path, she must consider a total of 5040 paths! ■

Example.

How many permutations of the letters ABCDEFGH contain the string ABC ?

Solution.

Because the letters ABC must occur as a block, we can find the answer by finding the number of permutations of six objects, namely, the block ABC and the individual letters D, E, F , G, and H. Because these six objects can occur in any order, there are 6! = 720 permutations of the letters ABCDEFGH in which ABC occurs as a block. ■

●Combinations

● We now turn our attention to counting unordered selection of objects.

Example.

How many different committees of three students can be formed from a group of four students?

Solution.

We need only find the number of subsets with three elements from the set containing the four students. We see that there are four such subsets, one for each of the four students, because choosing three students is the same as choosing one of the four students to leave out of the group. This means that there are four ways to choose the three students for the committee, where the order in which these students are chosen does not matter. ■

Definition.

An ***r*-combination** of elements of a set is an unordered selection of *r* elements from the set.

Thus, an *r*-combination is simply a subset of the set with *r* elements.

Example.

Let $S = \{1, 2, 3, 4, 5\}$. Then $\{1, 3, 4\}$ is a 3- combination of S. (Note that $\{4, 1, 3\}$ is the same 3-combination as $\{1, 3, 4\}$, because the order in which the elements of a set are listed does not matter.)

Definition.

The number of *r*-combinations of a set with n elements, where n is a nonnegative integer and *r* is an integer with $0 \le r \le n$, equals

$$\boldsymbol{(n, r), nCr\ or\ \binom{n}{r} = \frac{n!}{r!(n-r)!}}.$$

Example.

We see that $\boldsymbol{C(4,2) = 6}$, because the 2-combinations of $\{\boldsymbol{a,b,c,d}\}$ are the six subsets $\{\boldsymbol{a,b}\}, \{\boldsymbol{a,c}\}, \{\boldsymbol{a,d}\}, \{\boldsymbol{b,c}\}$, $\{\boldsymbol{b,d}\}$, and $\{\boldsymbol{c,d}\}$. ■

Example.

How many poker hands of five cards can be dealt from a standard deck of 52 cards? Also, how many ways are there to select 47 cards from a standard deck of 52 cards?

Solution.

Because the order in which the five cards are dealt from a deck of 52 cards does not matter, there are

$\boldsymbol{C(52,5) = 2,598,960}$

different hands of five cards that can be dealt.

Consequently, there are 2,598,960 different poker hands of five cards that can be dealt from a standard deck of 52 cards.

Note that there are $\boldsymbol{C(52,47) = 2,598,960}$ different ways to select 47 cards from a standard deck of 52 cards.

■

Example.

How many ways are there to select five players from a 10-member tennis team to make a trip to a match at another school?

Solution.

The answer is given by the number of 5-combinations of a set with 10 elements. The number of such combinations is $C(10, 5) = 252$. ■

Example.

A group of 30 people have been trained as astronauts to go on the first mission to Mars. How many ways are there to select a crew of six people to go on this mission (assuming that all crew members have the same job)?

Solution

The number of ways to select a crew of six from the pool of 30 people is the number of 6-combinations of a set with 30 elements, because the order in which these people are chosen does not matter. The number of such combinations is $C(30, 6) = 593{,}775$. ■

Example.

How many bit strings of length *n* contain exactly *r* 1s?

Solution.

The positions of *r* 1s in a bit string of length *n* form an *r*-combination of the set $\{1, 2, 3, \ldots, n\}$. Hence, there are $C(n, r)$ bit strings of length *n* that contain exactly *r* 1s.

Example.

Suppose that there are 9 faculty members in the mathematics department and 11 in the computer science department. How many ways are there to select a committee to develop a discrete mathematics course at a school if the committee is to consist of three faculty members from the mathematics department and four from the computer science department?

Solution.

By the product rule, the answer is the product of the number of 3-combinations of a set with nine elements and the number of 4-combinations of a set with 11 elements. By Theorem 2, the number of ways to select the committee is $\boldsymbol{C(9,3) \cdot C(11,4) = 27,720}$.■

♣ Combinatorial Proofs

To remind, $C(n, m)$ is a *binomial coefficient*

$$\mathrm{C}(\mathrm{n}, \mathrm{m}) = \frac{\mathrm{n}!}{\mathrm{m}!\,(\mathrm{n}-\mathrm{m})!}$$

that appears in the Binomial Theorem which, for an integer exponent, can be written as

$(x + y)^n$

$= C(n, 0)\, x^n + C(n, 1)\, x^{n-1} y + C(n, 2)\, x^{n-2} y^2 + ... + C(n, n)\, y^n$

$= \sum_{k=0}^{n} C(n, k) x^{n-k} y^k.$

● Combinatorial proof is a perfect way of establishing certain algebraic identities without resorting to any kind of algebra. For example, let's consider the simplest property of the ***binomial coefficients***:

(1) $C(n,\ k)\ =\ C(n, n\ -\ k).$

To prove this identity we do not need the actual ***algebraic formula*** that involves factorials, although this, too, would be simple enough. All that is needed to prove (1) is the knowledge of ***the definition***: $C(n, k)$ denotes the number of ways to select k out n objects without regard for the order in which they are selected. To prove (1) one needs to observe that whenever k items are selected, *n*-*k* items are left over, (un)selected of sorts. So that proving (1) becomes a word usage matter. (In this example, another simple proof is by introducing $m = n - k$, from which $k = n - m$ so that (1) translates into an equivalent form $C(n, n - m) = C(n, m)$.).◄

As another example, the identity

(2) $C(n, 0) + C(n, 1) + C(n, 2) + \ldots + (n, n - 1), + C(n, n) = 2^n$

which is a consequence of the ***binomial theorem***

$(x + y)^n = \Sigma$ C(n, k) x^k y^{n-k}, $0 \leq k \leq n$.

admits a combinatorial interpretation. The left hand side in (2) represents the number of ways to select a group - empty or not - of items out of a set of *n* distinct elements. The first term gives the number of ways not to make any selection, which is 1. The second term gives the number of ways to select one item (which is *n*), etc. What does the right hand side represent? Exactly same thing. Indeed, with every selection of items from a given set we can associate a ***function*** that takes values 0 or 1. A selected element is assigned value 1, while an unselected element is assigned value 0. If for the sake of counting convenience, the elements of the set are ordered with indices 1, ..., *n*, then every selection from the set is represented by a string of 0's and 1's; the total number of such strings is clearly the right hand side in (2): 2^n.◄

Thus a combinatorial proof consists in providing two answers to the same question. But not to forget, finding the question to be answered in two ways is conceivably the most important part of the proof. As a matter of convention, it is often convenient to think of sets and their elements as groups of students and of selections of elements as endowing them with a membership on a committee. For a third example, consider the popular identity underlying the ***Pascal triangle***:

(3) $C(n,k) = C(n-1,k) + C(n-1,k-1)$.

By definition, the left hand side is the number of ways to compose a *k*-member committee out of a group of *n* students. To grasp the significance of the right hand side, pick arbitrarily one of the students. Then the first term on the right gives the number of *k*-member committees that do not include the student, whereas the second term gives the number of committees in which the student is a member.◄

Here is an additional example. Prove that

(4) $C(n,r)\,C(r,k) = C(n,k)\,C(n-k,r-k)$,

where $k \leq r \leq n$. *C*(*n*, *r*) is the number of ways to form an *r*-member committee from a group of *n* students. *C*(*r*, *k*) is the number of ways to form a *k*-member committee out of a group of *r* students. As *r* is the same in both cases, it it sensible to assume that the *r* students selected from the initial *n* are exactly those among whom we seek a more restrictive *r*. So we could describe the left hand side in (4) as the number of ways to choose a *k*-member committee from an *n*-member student body and a *k*-member subcommittee out of the selected *r*. So the question is in how many ways is it possible to choose a *r*-member committee from an *n*-member student body and a *k*-member subcommittee out of the selected *r*. The left hand side gives an answer to that question. The right hand side answers the same question but in a different way. First we select a *k*-member subcommittee out of the *n*-member student population and later complete it to an *r*-

member committee by selecting r-k members out of the remaining population of n-k students.◄

Example.

How many different strings can be made by reordering the letters of the word TOTTOS?

Solution.

If all letters in the word TOTTOS would be different, then the answer would be 5! ! but then we would over count. To avoid it, we observe that there are 6 positions. The letter T can be placed among these six positions in C(6, 3) times, while the letter O can be placed in the remaining positions in C(3, 2) ways; finally S can be put in C(1, 1) ways. By the multiplication rule we have

C(6, 3) C(3, 2) C(1, 1) = 60 orderings.■

Exercise (6)

1-Cells of a 15×15 square grid have been painted in red, blue and green. Prove that there are at least two rows of cells with the same number of squares of at least one of the colors.

2-There are $(2n - 1)$ rooks on a $(2n - 1) \times (2n - 1)$ board placed so that none of them threatens another. Prove that any $n \times n$ square contains at least one rook.

3- In every square of a 5×5 board there is a flea. At some point, all the fleas jump to an adjacent square (two squares are adjacent if they share an edge). Is it possible that after they settle in the new squares, the configuration is exactly as before: one flea per square?

4- 200 points have been chosen on a circle, all with integer number of degrees. Prove that the points there are at least one pair of antipodes, i.e., the points 180° apart.

5- If each point of the plane is colored red or blue then there are two points of the same color at distance 1 from each other.

6-The integers 1, 2, ..., 10 are written on a circle, in any order. Show that there are 3 adjacent numbers whose sum is 17 or greater.

7-Given a planar set of 25 points such that among any three of them there exists a pair at the distance less than 1. Prove that there exists a circle of radius 1 that contains at least 13 of the given points.

8-Prove that among any five points selected inside an equilateral triangle with side equal to 1, there always exists a pair at the distance not greater than .5.

9-Let *A* be any set of 19 distinct integers chosen from the arithmetic progression 1, 4, 7,..., 100. Prove that there must be two distinct integers in A whose sum is 104.

10-Prove that in any set of 51 points inside a unit square, there are always three points that can be covered by a circle of radius 1/7.

11-Five points are chosen at the nodes of a square lattice (grid). Why is it certain that at least one mid-point of a line joining a pair of chosen points, is also a lattice point?

12-Prove that there exist two powers of 3 whose difference is divisible by 1997.

13-If 9 people are seated in a row of 12 chairs, then some consecutive set of 3 chairs are filled with people.

14-Given any sequence of n integers, positive or negative, not necessarily all different, some consecutive subsequence has the property that the sum of the members of the subsequence is a multiple of n.

15-In every polyhedron there is at least one pair of faces with the same number of sides.

16-In every polyhedron there is at least one pair of faces with the same number of sides.

17-Given 12 distinct 2-digit integers. Prove there are some two whose difference - a 2-digit number - has equal digits.

18-What is the largest number of cells of a 6×6 board that could be colored such that no two colored cells touch (not even at a corner)?

19-17 students talked of 3 topics. There are 3 students that - between them - talked the same topic.

20-Seven integers under 127 and their Ratios

21-17 rooks are placed on an 8×8 chessboard. Prove that there are at least 3 rooks that do not threaten each other.

22-Chinese Remainder Theorem.

Let's mark the centers of all squares of an 8x8 chess board. Is it possible to cut the board with 13 straight lines (none passing through a single midpoint) so that every piece had at most 1 marked point?

23-Each of the given 9 lines cuts a given square into two quadrilaterals whose areas are in proportion 2:3. Prove that at least three of these lines pass through the same point.

24-Suppose each point of the plane is colored red or blue. Show that some rectangle has its vertices all the same color.

25-Suppose each point on a circumference of a circle is colored either red or blue. Prove that, no matter how colors may be distributed, there exist 3 *equally spaced* points of the same color.

26-Suppose $f(x)$ is a polynomial with integral coefficients. If $f(x) = 2$ for three different integers a, b, and c, prove that, for no integer, $f(x)$ can be equal to 3.

27-Prove that there exists a power of three that ends with 001.

28-Show that if more than half of the subsets of an n-element set are selected, then some two of the selected subsets have the property that one is a subset of the other.

29-Let *a* and *b* be positive integers, with $a < b < 2a$. Then, given more than half of the integers in the set $\{1, 2, \ldots, a + b\}$, some two of the given integers differ by a or by b.

30-Given any 6 points inside a circle of radius 1, some two of the 6 points are within 1 of each other.

31-Let *n* be a positive integer greater than 3. Let *m* be the largest integer in $(n + 2)/2$. Then, given more than *m* of the integers in the set $\{1, 2, \ldots, n\}$, some three of the integers in the given set have the property that one of the three is the sum of the other two.

32-If more than half of the integers from $\{1, 2, \ldots, 2n\}$ are selected, then some two of the selected integers have the property that one divides the other.

33-If more than half of the integers from $\{1, 2, \ldots, 2n\}$ are selected, then some two of the selected integers are mutually prime.

34-Given any sequence of $mn + 1$ real numbers, some subsequence of $(m + 1)$ numbers is increasing or some subsequence of $(n + 1)$ numbers is decreasing.

35-Given any 1000 integers, some two of them differ by, or sum to, a multiple of 1997.

36-Given any 10 4-element subsets of an 11-set, some two of the subsets intersect in at least two elements.

37-A person takes at least one aspirin a day for 30 days. If he takes 45 aspirin altogether, in some sequence of consecutive days he takes exactly 14 aspirin.

38-A theatre club gives 7 plays one season. Five women in the club are each cast in 3 of the plays. Then some play has at least 3 women in its cast.

39-At a party of n people, some pair of people are friends with the same number of people at the party.

40-Given any 6 integers from 1 to 10, some two of them have an odd sum.

41. How many ways are there to select a first-prize winner, a second-prize winner, and a third-prize winner from 100 different people who have entered a contest?

42. How many permutations of the letters ABCDEFG contain the string ABC?

43. How many different committees of two students can be formed from a group of four students?

Answer $C(4,2)$.

44. In how many ways can we choose a chair and a vice chair from a group of four students?

How is this example different from the previous one?

45. How many poker hands of five cards can be dealt from a standard deck of 52 cards? Also, how many ways are there to select 47 cards from a standard deck of 52 cards?

46. How many bit strings of length *n* contain exactly *r* 1s?

Chapter (VII)

Boolean Algebra and Their Applications

Chapter (VII)

Boolean Algebra and Their Applications

The circuits in computers and other electronic devices have inputs, each of which is either a 0 or a 1, and produce outputs that are also 0s and 1s. Circuits can be constructed using any basic element that has two different states. Such elements include switches that can be in either the on or the off position and optical devices that can either be lit or unlit. In 1938 Claude Shannon showed how the basic rules of logic, first given by George Boole in 1854 in his book “The Laws of Thought” could be used to design circuits. These rules form the basis for Boolean algebra.

7.1 Boolean Algebra

Definition.

The Boolean algebra is a mathematical system $(B, +, \cdot, \acute{\square}, 0, 1)$, where B is a non-empty set (contains at least two elements $\{0, 1\}$) , $+, \cdot$ are two binary operations on B i .e. , $" + ", " \cdot "$ are maps from $B \times B \rightarrow B$), $"\acute{\square}"$ is a unary operation, i.e., $\acute{\square}: B \rightarrow B$; and $0 \neq 1$ two elements in B, satisfies for all $a, b, c \in B,$ the following axioms:

1. There exist at least two elements a, b in B and that $a \neq b.$

2. $\forall\ a, b \in B$

(i) $a + b \in B$, (ii) $a \cdot b \in B$;

3. Commutative laws: for all $a, b \in B$,

(i) $a + b = b + a,$ (ii) $a \cdot b = b \cdot a$;

4. Associative laws: for all $a, b, c \in B$,

(i) $a + (b + c) = (a + b) + c$,

(ii) $a \cdot (b \cdot c) = (a \cdot b) \cdot c$;

5. Distributive laws: $a, b, c \in B$,

(i) $a \cdot (b + c) = a \cdot b + a \cdot c,$

(ii) $a + (b \cdot c) = (a + b) \cdot (a + c)$;

6. (i) Existence of zero: There exists of B such that

$a + 0 = a\ \forall a \in B$;

The element 0 is called the zero element.

(ii) Existence of identity (unit): There exists $1 \ \in \ B$ such that

$a \cdot 1 = a\ \forall a \in B$;

The element 1 is called the identity (unit) element.

7. Existence of complement:

For each $a \in B, \exists a' \in B$ such that

(i) $a + a' = 1$, (ii) $a \cdot a' = 0$.

Example.

Let $B_2 = \{0,1\}$ and $+, \cdot, \ '$ be defined as follows :

a	a'
1	0
0	1

a	b	$a+b$	$a \ \cdot b$
1	1	1	1
1	0	1	0
0	1	1	0
0	0	0	0

Then $(B, +, \cdot, \square', 0, 1)$ is a Boolean algebra. ■

Example.

Let X be a non-empty set, then $\big(P(X), +, \cdot, \square', 0, 1\big)$, is a Boolean algebra, where for all $A, B \in P(X)$

$A + B = A \cup B, A \cdot B = A \cap B, \ A' = A^c, 0 = \phi, 1 = X.$■

Example.

Suppose B is the set of all propositions, then $\big(B, +, \cdot, \square', 0, 1\big)$ is a Boolean algebra, where $p + q = p \vee q, p \cdot q = p \wedge q, p' = \sim p, 0 = F, 1 = T$, for all $p, q \in B$, and T is the tautology and F is the contradiction. ■

Example.

If B is the set of all positive divisors for 30 , then the system $(B,+,\cdot,\acute{\square},0,1)$ is a Boolean Algebra, where $x+y=\text{lcm}[x,y], x.y=\gcd(x,y), x'=\dfrac{30}{x}, 0=1,$ $1=30$. ■

Example.

Let B be the set of all positive divisors for 8, then $(B,+,\cdot,\acute{\square},0,1)$ is not Boolean Algebra, where the operations is defined in the above example. Because $4+4'=\text{lcm}[4,2]=4\neq 8=1.$ ■

Example.

Let S be the set of statement formulas involving n statement variables. The algebraic system $(S,\ \vee,\ \wedge,\neg,\ F,\ T)$ is a Boolean algebra in which $\vee,\ \wedge,\neg$ denotes the operations of conjunction, disjunction and negation, respectively. The element F and T denotes the formulas which are contradictions and Tautologies, respectively. ■

Remark.

We use the symbol B instead of the Boolean Algebra $(B, +, \cdot, \acute{\square}, 0, 1)$ and use ab instead of $a \cdot b$.

Theorem.

Let B be a Boolean Algebra. Then

(1) There is at most one identity element w. r. t."$+$", i. e., the additive identity 0 is unique.

(2) There is at most one identity element w. r. t."$\cdot$", i. e., the multiplicative identity 1 is unique.

(3) The complement a' of a is unique.

Proof.

(1) Let $0'$ be another additive identity.

Since $a = a + 0'$, then $0 = 0 + 0' = 0' + 0 = 0'$.

(2) Let $1'$ be another multiplicative identity.

Then $1 = 1 \cdot 1' = 1' \cdot 1 = 1'$.

(3) Let $y \in B$ be another complement of a,

i. e., $a + y = 1$ and $a \cdot y = 0$.

Then

$y = y \cdot 1$ (Identity element)

$= y(a + a')$ (Complemented)

$= ya + ya'$ (Distributive law)

$= ay + ya'$ (Commutative law)

$= 0 + ya'$ (Complemented)

$= aa' + ya'$ (Complemented)

$= (a + y)a'$ (Distributive law)

$= 1 \cdot a'$ (Complemented)

$= a'.$ (Identity element)

Therefore a' is the unique complement of a . ■

Theorem. **Double Complement law**

For every element a in a Boolean algebra B, $(a')' = a$.

Proof:

Suppose B is a Boolean algebra and a is any element of B. Then

$a + a' = a' + a$ by the complement law for $+$.

$= 1$ by the complement law for 1. and

$a \cdot a' = a' \cdot a$ by the complement law for $\cdot$.

$= 0$ by the complement law for 0.

Thus a satisfies the two equations with respect to a' that are satisfied by the complement of a'. From the fact that the complement of a is unique,

we conclude that $(a')' = a$.◄

Example.

Fill in the blanks in the following proof that for all elements a in a Boolean algebra B, $a + a = a$.

Proof.

Suppose B is a Boolean algebra and a is any element of B. Then

$a = a + 0$ (a)

$= a + a \cdot a'$ (b)

$= (a + a) \cdot (a + a')$ (c)

$= (a + a) \cdot 1$ (d)

$= a + a$ (e)

Solution.

(a) because 0 is an identity for1

(b) by the complement law for ·

(c) by the distributive law for1over?

(d) by the complement law for1

(e) because 1 is an identity for? ■

Definition.

The Boolean expression in the variables $x_1, x_2, \dots, x_n$ are defined recursively as $0,1, x_1, x_2, \dots, x_n$.

If E and F are Boolean expression , then $E', E + F, E.F$ are Boolean expression.

Definition.

The dual of a Boolean expression E is denoted by E^d and is obtained by interchanging Boolean Sum "+" and Boolean products "."; and interchanging 0_s and 1_s.

Example.

Find the duals of

(1) $E = x(y + 0)$;

(2) $T = xz' + x \cdot 0 + x' \cdot 1$;

(3) $F = x' \cdot 1 + (y' + z)$;

Solution.

(1) $E^d = x + y \cdot 1$.

(2) $T^d = (x + z') \cdot (x + 1) \cdot (x' + 0)$.

(3) $F^d = (x' + 0) \cdot (y' \cdot z)$.■

Theorem. (Identities (Rules) of Boolean Algebra)

Let B be Boolean algebra and $a, b \in B$. Then

(1) Idempotent rules:

(a) $a + a = a,$ (b) $aa = a;$

(2) Identity rules:

(a) $a + 1 = 1\ ,$ (b) $a0 = 0;$

(3) Absorption rules:

(a) $a + ab = a,$ (b) $a(a + b) = a;$

(4) (a) $0' = 1,$ (b) $1' = 0;$

(5) De Morgan's rules:

(a) $(a + b)' = a'b'\ ,$ (b) $(ab)' = a' + b'.$

Proof. All the given properties are Boolean expression and it's dual. So, by the duality principle we only prove one of these expressions.

(1) (a) See the previous example.

(2) (a) $a + 1 = (a + 1)1$ Identity

$= (a + 1)(a + a')$ Identity

$= a + (1a')$ Distributive

$= a + (a'.1)$ Commutative

$= a + a'$ Identity

$= 1.$ Complement

(3) (a) $a+ab=a1+ab$ Identity

$$\begin{aligned} &= a(1+b) && \text{Distributive}\\ &= a(b+1) && \text{Commutative}\\ &= a1 && \text{From (2) above}\\ &= a. && \text{Identity}\end{aligned}$$

(4) Since $0+1=1$ and $0.1=0$, then by uniqueness of the complemented we obtain $0'=1$ and $1'\ =\ 0$.

(5) $(a+b)(a'b')=(a'b')(a+b)$ Commutative

$$\begin{aligned} &= (a'b')a+(a'b')b && \text{Distributive}\\ &= a(a'b')+(a'b')b && \text{Commutative}\\ &= (aa')b'+a'(bb') && \text{Associative}\\ &= 0b'+a'0 && \text{Identity}\\ &= 0+0 && \text{From (2) above}\\ &= 0. && \text{Zero}\end{aligned}$$

$$\begin{aligned}(a+b)+a'b' &= \big((a+b)+a'\big)\big((a+b)+b'\big)\\ &= \big((b+a)+a'\big)\big((a+b)+b'\big)\\ &= \big(b+(a+a')\big)\big(a+(b+b')\big)\\ &= (b+1)(a+1)=1\cdot 1=1.\end{aligned}$$

By uniqueness of the complement, we have

$a'b'=(a+b)'$.◀

Also, we can prove the identities in the above theorem by the truth table as shown in the following example.

Example.

Show that the distributive law

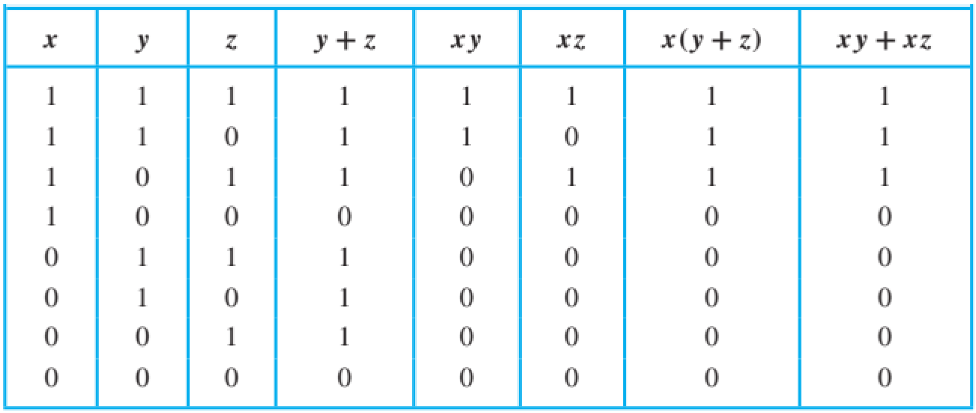
$$x(y+z) = xy + xz$$

is valid.

Solution.

The verification of this identity is shown in the following table, where $1+1=1\,,1+0=1\,,0+1=1,\ 0+0=0\,$, and $1.1=1\,,1.0=0\,,0.1=0\,,0.0=0.$

The identity holds because the last two columns of the table agree.



x	y	z	$y+z$	xy	xz	$x(y+z)$	$xy+xz$
1	1	1	1	1	1	1	1
1	1	0	1	1	0	1	1
1	0	1	1	0	1	1	1
1	0	0	0	0	0	0	0
0	1	1	1	0	0	0	0
0	1	0	1	0	0	0	0
0	0	1	1	0	0	0	0
0	0	0	0	0	0	0	0

Theorem. [Duality principle]

If T is an identity in a Boolean algebra B, then T^d is also an identity in B.

This result is useful for obtaining new identities.

Example.

By taking duals construct an identity form the following absorption law:

$$x(x + y) = x.$$

Solution.

Taking the duals of both sides of this identity produces the identity $x + xy = x$, which is also called an absorption law. ■

Example.

Find the duals of $x(y\ +\ 0)$ and $x' \cdot 1 + (y' + z)$.

Solution: Interchanging $\cdot$ signs and $+$ signs and interchanging 0s and 1s in these expressions produces their duals. The duals are $x\ +\ (y\ \cdot\ 1)$ and $(x' + 0)(y'z)$, respectively. ■

7.2 Boolean Functions

Definition.

Let $n \in Z^+$ and $B_2^n = \{(x_1, x_2, \dots, x_n); x_i \in B_2\}$ be the set of all possible n- tuples of 0_s and 1_s. Then the function $f: B_2^n \to B_2$ is said to be Boolean function of degree n and n variables $x_1, \dots, x_n$.

We will use examples to illustrate one important way to find a Boolean expression that represents a Boolean function.

Example.

Find Boolean expressions that represent the functions $F(x, y, z)$ and $G(x, y, z)$, which are given in the Table 1.

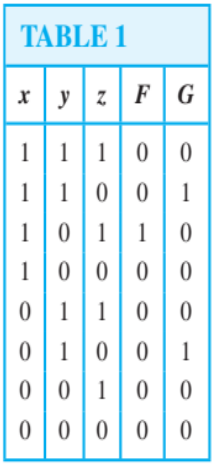


TABLE 1

x	y	z	F	G
1	1	1	0	0
1	1	0	0	1
1	0	1	1	0
1	0	0	0	0
0	1	1	0	0
0	1	0	0	1
0	0	1	0	0
0	0	0	0	0

Solution.

●An expression that has the value 1 when $x = z = 1$ and $y = 0$, and the value 0 otherwise, is needed to represent F.

●Such an expression can be formed by taking the Boolean product of x, y', and z.

●This product, $xy'z$, has the value 1 if and only if

$x = y' = z = 1$, which holds if and only if $x = z = 1$ and $y = 0$.

●To represent G, we need an expression that equals 1 when $x = y = 1$ and $z = 0$, or $x = z = 0$ and $y = 1$.

●We can form an expression with these values by taking the Boolean sum of two different Boolean products.

●The Boolean product xyz' has the value 1 if and only if $x = y = 1$ and $z = 0$. Similarly, the product $x'yz'$ has the value 1 if and only if $x = z = 0$ and $y = 1$.

●The Boolean sum of these two products, $xyz' + x'yz'$, represents G, because it has the value 1 if and only if $x = \ y = 1$ and $z = 0$, or $x = z = 0$ and $y = 1$. ■

To represent the Boolean function we use the truth table for it.

Example.

Find the values of the Boolean function represented by

$$F(x, y) = xy + x'.$$

Solution.

The values of this function are displayed in the following table

x	y	xy	x'	$xy + x'$
1	1	1	0	1
1	0	0	0	0
0	1	0	1	1
0	0	0	1	1

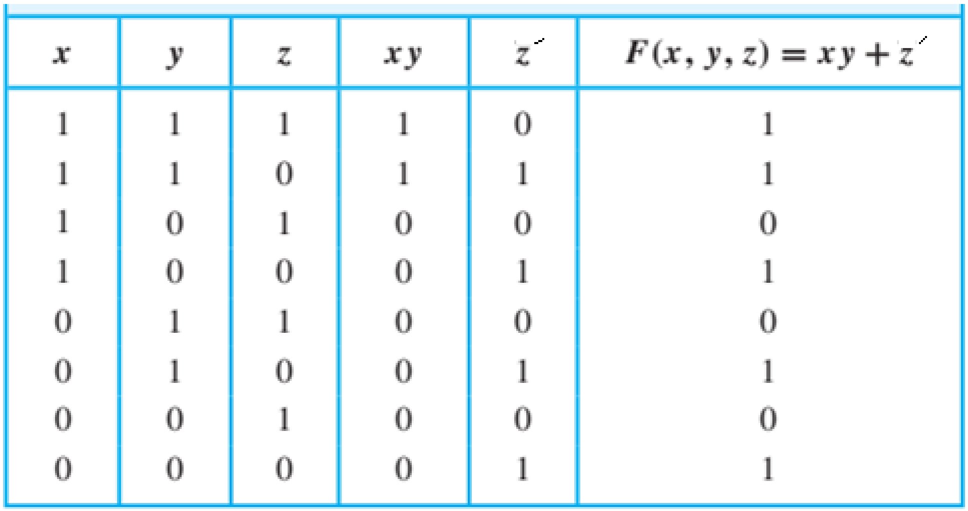
Example.

Find the values of the Boolean function represented by

$$F(x, y, z) = xy + z'.$$

Solution.

The values of this function are displayed in the following table

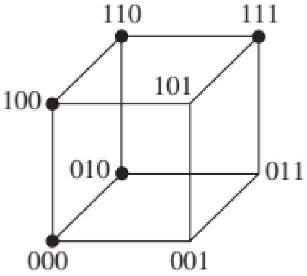


x	y	z	xy	z'	$F(x, y, z) = xy + z'$
1	1	1	1	0	1
1	1	0	1	1	1
1	0	1	0	0	0
1	0	0	0	1	1
0	1	1	0	0	0
0	1	0	0	1	1
0	0	1	0	0	0
0	0	0	0	1	1

Note that we can represent a Boolean function graphically by distinguishing the vertices of the *n*-cube that correspond to the *n*-tuples of bits where the function has value 1.

Example.

The function $F(x, y, z) = xy + z'$ from B_2^3 to B_2 from the above example can be represented by distinguishing the vertices that correspond to the five 3-tuples $(1, 1, 1)$, $(1, 1, 0)$, $(1, 0, 0)$, $(0, 1, 0)$, and $(0, 0, 0)$, where $F\ (x, y, z)\ =\ 1$, as shown in the given figure. These vertices are displayed using solid black circles. ■



Definition.

Let $f, g: B_2^n \rightarrow \mathrm{B}_2$ be Boolean Function then *f* and g are equivalent written $f \equiv g$ if and only if they have the same truth table or we can obtain one of them from the other.

Example.

Prove that the Boolean functions $f(x, y, z) = xy$ and $g(x, y, z) = xy(xz' + yz)$ are equivalent.

Solution.

$$\begin{aligned} g(x,y,z) &= xy(xz' + yz) \\ &= (xy)(xz') + (xy)(yz) \\ &= (yx)(xz') + x(yy)z \\ &= y(xx)z' + x(yy)z \\ &= yxz' + xyz \\ &= (xy)z' + (xy)z \\ &= xy(z' + z) \\ &= xy \cdot 1 = xy. \end{aligned}$$

We leave the student to prove it using truth table. ■

Definition.

Let f and g be two Boolean functions in n variables. We define the Boolean sum $f + g$, Boolean product $f \cdot g$ and f' as follows:

$(f + g)(x_1, \dots, x_n) = f(x_1, \dots, x_n) + g(x_1, \dots, x_n);$

$(f \cdot g)(x_1, \dots, x_n) = f(x_1, \dots, x_n). g(x_1, \dots, x_n);$

$f'(x_1, \dots, x_n) = [f(x_1, \dots, x_n)]'.$

Remark.

The algebraic system $(F_n, +, \cdot, \square', 0,1)$ is a Boolean algebra where $n \in \mathbb{Z}^+$, F_n is the set of all Boolean function, $0(x_1, \dots, x_n) = 0$ and $1(x_1, \dots, x_n) = 1$

Definition.

Let $f(x_1, \dots, x_n)$ be a Boolean function.

(a) For every $1 \le i \le n$, x_i or $x_i{}'$ is called **literal**.

(b) The product $y_1 y_2 \dots y_{n,}$ $y_i = x_i$ or $y_i = x_i{}'$ for every $1 \le i \le n$ is called **minterm**.

Hence, a minterm is a product of n literals, with one literal for each variable. A minterm has the value 1 for one and only one combination of values of its variables. More precisely, the minterm $y_1 y_2 \dots y_n$ is 1 if and only if each y_i is 1 and this occurs if and only if $x_i = 1$ when $y_i = x_i$ and $x_i = 0$ when $y_i = x_i{}'$.

Example.

Find a minterm that equals 1 if $x_1 = x_3 = 0$ and $x_2 = x_4 = x_5$ =1 and equals 0 otherwise.

Solution.

The minterm $x_1{}' x_2 x_3{}' x_4 x_5$ has the correct set of values.

Definition.

If f is written as the sum of minterm , then f is in the **Complete Sum of Products** (**CSP**) or **Sum – of – Products Expansion.**

We can put any function in CSP as follows:

(1) Find the truth table of f.

(2) Determine the rows that have the value 1.

(3) Find the minterm $y_1 y_2 \dots . y_n$ for each row in step (2), where $y_i = x_i$ if $x_i = 1$ and $y_i = x_i{}'$ if the value of x_i is 0.

(4) CSP of f is the sum of the minterms obtained in (3).

Definition.

Using the duality principle we can obtain the **Complete Product of Sums (CPS) or Product – of – Sums Expansion** of a Boolean function as follows:

1. Find the truth table of f.

2. Determine the rows that have the value 0.

3. Find the **maxterm** $y_1 + y_2 + \cdots + y_n$ for each row in 2, where $y_i = x_i$ if $x_i = 0$ and $y_i = x_i{}'$ if $x_i = 1$.

4. CPS of f is the product of maxterms obtained in 3.

Example.

Find CSP(F) and CPS(F),where $F(x, y, z) = (x + y)z'$.

Solution.

We can construct the CSP(F) and CPS(F) by determining the truth table:

x	y	z	$x + y$	z'	$F = (x+y)z'$	minterm	maxterm
1	1	1	1	0	0		$x' + y' + z'$
1	1	0	1	1	1	xyz'	
1	0	1	1	0	0		$x' + y + z'$
1	0	0	1	1	1	$xy'z'$	
0	1	1	1	0	0		$x + y' + z'$
0	1	0	1	1	1	$x'yz'$	
0	0	1	0	0	0		$x + y + z'$
0	0	0	0	1	0		$x + y + z$

From the above table we have that the minterms of F are $xyz', xy'z'$ and $x'yz'$. Therefore

$$\text{CSP}(F) = xyz' + xy'z' + x'yz'.$$

Also, from the above table we have that the maxterms of F are $x' + y' + z'$, $x' + y + z', x + y' + z'$, $x + y + z'$ and $x + y + z$.

Therefore

$\text{CPS}(F) = (x + y + z)(x + y + z')(x + y' + z')$
$(x' + y + z')(x' + y' + z')$.■

We can obtain CSP(F) by the properties of the Boolean algebra as follows:

$F(x, y, z) = (x + y)z'$

$= xz' + yz'$ Distributive law

$= x1z' + 1yz'$ Identity law

$$= x(y+y')z' + (x+x')yz' \quad \text{Unit property}$$

$$= xyz' + xy'z' + xyz' + x'yz' \quad \text{Distributive law}$$

$$= xyz' + xy'z' + x'yz' \quad \text{Idempotent law}$$

So, CSP(F) = $xyz' + xy'z' + x'yz'$.

Also, We can obtain CPS(F) by the properties of the Boolean algebra as follows:

$$F(x,y,z) = (x+y)z'$$

$$= (x+y+0)(0+z')$$

$$= \big(x+y+(zz')\big)\big((xx')+z'\big)$$

$$= (x+y+z)(x+y+z')(x+z')(x'+z')$$

$$= (x+y+z)(x+y+z')(x+0+z')(x'+0+z')$$

$$= (x+y+z)(x+y+z')(x+yy'+z')(x'+yy' + z')$$

$$= (x+y+z)(x+y+z')(x+yy'+z')(x'+yy' + z')$$

$$= (x+y+z)(x+y+z')(x+y+z')(x+y'+z')(x' + y+z')(x'+y'+z')$$

$$= (x+y+z)(x+y+z')(x+y'+z')(x'+y + z')(x'+y'+z')$$

Note

● The CSP(F) is unique (except ordering of the minterms).

● The CPS(F) is unique (except ordering of maxterms).

●We can obtained CPS by giving the CSP for the complement of the function, and we take the complement of the CSP give the CPS.

$$\mathrm{CPS}(F) \;=\; \big(\mathrm{CSP}(F')\big)'$$

●If CSP$(F') = m_1 + m_2 + \cdots + m_k$, then CPS$(F) = \mathrm{m_1}'\mathrm{m_2}' \ldots \mathrm{m_k}'$.

So, we can obtain CPS(F) from CSP(F) as illustrated in the following example:

Example.

Use CSP (F') to find CPS (F) for the Boolean function $f(x, y, z) = xy + xz'$.

Solution.

We use the identities of the Boolean Algebra to find CPS(F) by obtaining CSP(F') algebraically:

$F(x, y, z) = xy + x'z$

$F'(x, y, z) = (xy + x'z)'$

$= (xy)'(x'z)'$ De Morgan's

$= (x' + y')((x')' + z')$ De Morgan's

$= (x' + y')(x + z')$

$= x'x + x'z' + xy' + y'z'$

$= 0 + x'(y + y')z' + xy'(z + z') + (x + x')y'z'$

$= x'yz' + x'y'z' + xy'z + xy'z' + xy'z' + x'y'z'$

$= x'yz' + x'y'z' + xy'z + xy'z'$

Therefore

$\mathrm{CSP}(F') = x'yz' + x'y'z' + xy'z + xy'z'$

and

$\mathrm{CPS}(F) = (x + y' + z)(x + y + z)(x' + y + z')$

$(x' + y + z)$.■

7.3 Logic Gates

●Boolean algebra is used to model the circuitry of electronic devices.

●Each input and each output of such a device can be thought of as a member of the set $\{0, 1\}$.

●A computer, or other electronic device, is made up of a number of circuits.

●Each circuit can be designed using the rules of Boolean algebra.

●The basic elements of circuits are called gates.

The three main ways of specifying the function of a combinational logic circuit are:

●Boolean Algebra. This forms the algebraic expression showing the operation of the logic circuit for each input variable either True or False that results in a logic "1" output.

●Truth Table. A truth table defines the function of a logic gate by providing a concise list that shows all the output states in tabular form for each possible combination of input variable that the gate could encounter.

●Logic Diagram. This is a graphical representation of a logic circuit that shows the wiring and connections of each individual logic gate, represented by a specific graphical symbol, that implements the logic circuit.

●The **inverter**, which accepts the value of one Boolean variable as input and produces the complement of this value as its output.

●The **OR gate**. The inputs to this gate are the values of two or more Boolean variables. The output is the Boolean sum of their values.

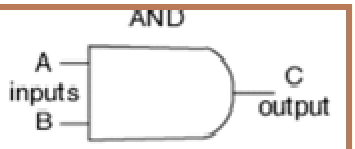
●The **AND gate**. The inputs to this gate are the values of two or more Boolean variables. The output is the Boolean product of their values.

●The **NAND gate** function is a combination of the two separate logical functions, the AND function and the NOT function in series.

● The **NOR gate** is also a combination of two separate logic functions, Not and OR connected together to form a single logic function which is the same as the OR function except that the output is inverted.

- AND gate

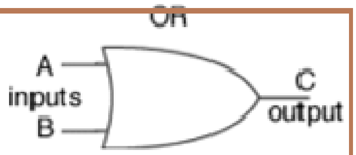
→ C = A • B



A	B	C
0	0	0
1	0	0
0	1	0
1	1	1

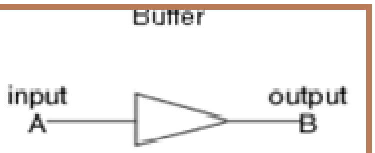
- OR gate

→ C = A+B



A	B	C
0	0	0
1	0	1
0	1	1
1	1	1

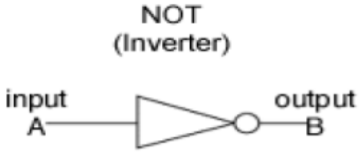
- Buffer



A	B
0	0
1	1

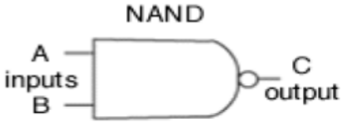
- NOT Gate

→ $B = A'$



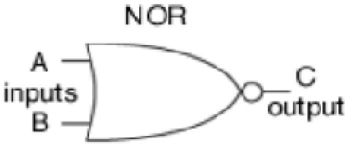
A	B
0	1
1	0

- NAND Gate



A	B	C
0	0	1
1	0	1
0	1	1
1	1	0

- NOR Gate



A	B	C
0	0	1
1	0	0
0	1	0
1	1	0

Combinations of Gates

Combinational circuits can be constructed using a combination of inverters, OR gates, and AND gates.

Example.

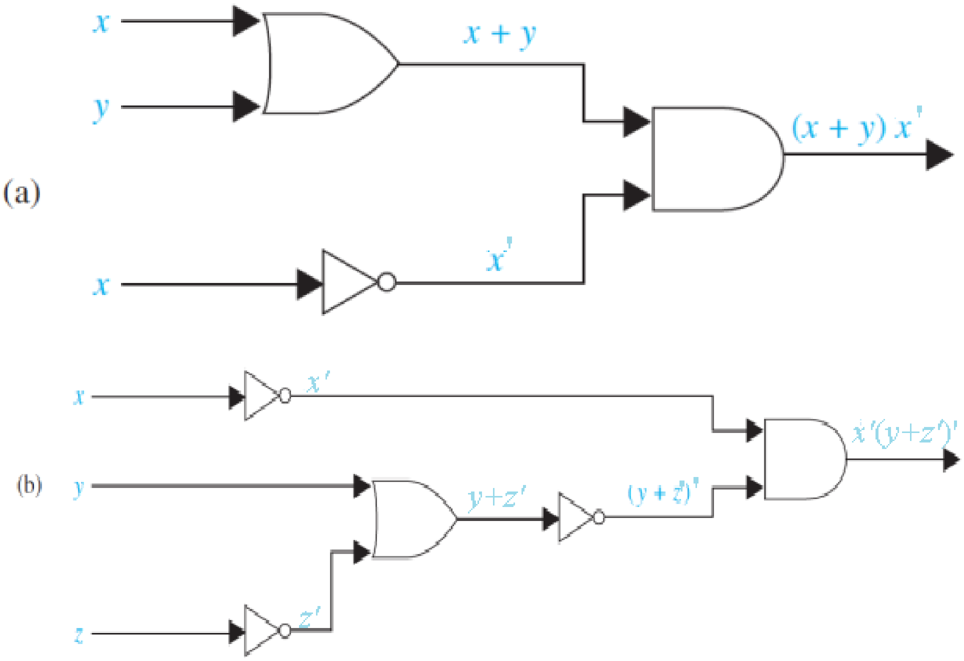
Construct circuits that produce the following outputs:

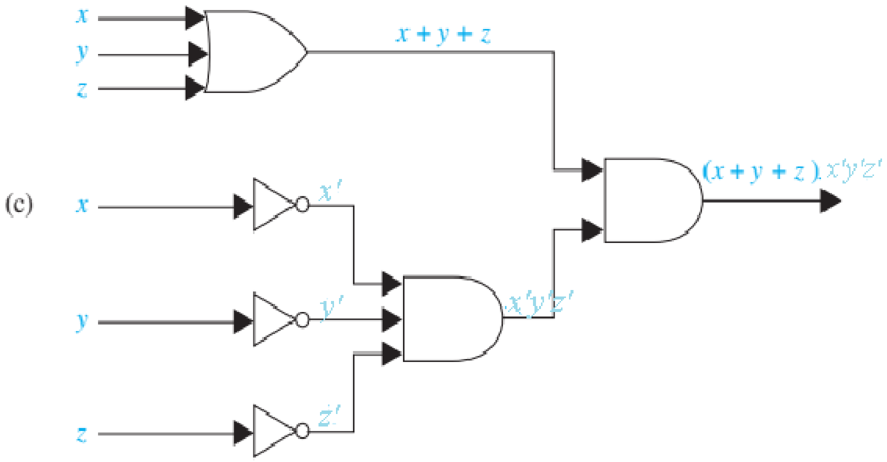
(a) $(x + y)x'$;

(b) $x'(y + z')'$;

(c) $(x + y + z)(x' y' z')$.

Solution.



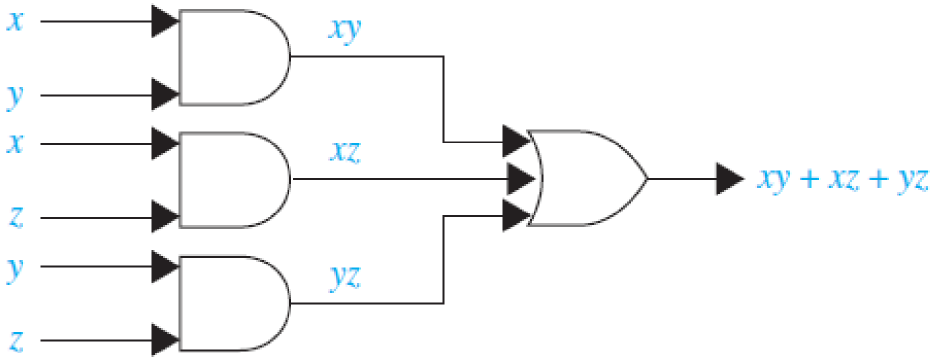


●Examples of Circuits

1. A committee of three individuals decides issues for an organization. Each individual votes either yes or no for each proposal that arises. A proposal is passed if it receives at least two yes votes. Design a circuit that determines whether a proposal passes.

Solution.

Let $x = 1$ if the first individual votes yes, and $x = 0$ if this individual votes no; let $y = 1$ if the second individual votes yes, and $y = 0$ if this individual votes no; let $z = 1$ if the third individual votes yes, and $z = 0$ if this individual votes no. Then a circuit must be designed that produces the output 1 from the inputs x, y, and z when two or more of x, y, and z are 1. One representation of the Boolean function that has these output values is $xy + xz + yz$. The circuit that implements this function is shown in the figure.



2. Sometimes light fixtures are controlled by more than one switch. Circuits need to be designed so that flipping any one of the switches for the fixture turns the light on when it is off and turns the light off when it is on. Design circuits that accomplish this when there are two switches and when there are three switches.

Solution.

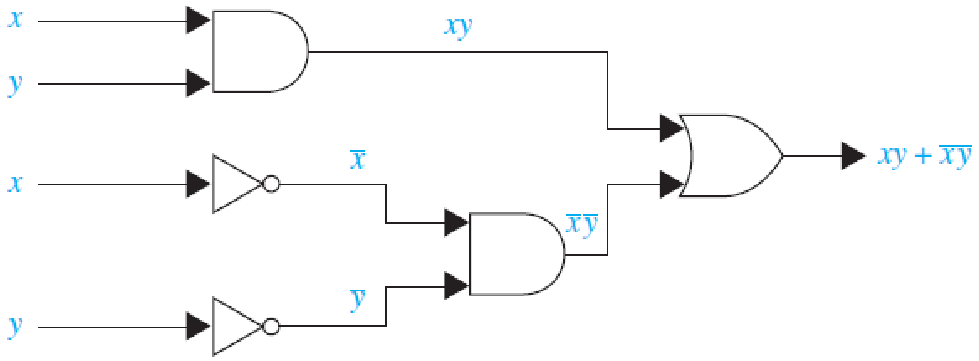
We will begin by designing the circuit that controls the light fixture when two different switches are used. Let $x = 1$ when the first switch is closed and $x = 0$ when it is open, and let $y = 1$ when the second switch is closed and $y = 0$ when it is open. Let $F(x, y) = 1$ when the light is on and $F(x, y) = 0$ when it is off. We can arbitrarily decide that the light will be on when both switches are closed, so that $F(1, 1) = 1$. This determines all the other values of F. When one of the two switches is opened, the light goes off, so $F(1, 0) = F(0, 1) = 0$. When the other switch is also opened, the light goes on, so $F(0, 0) = 1$.

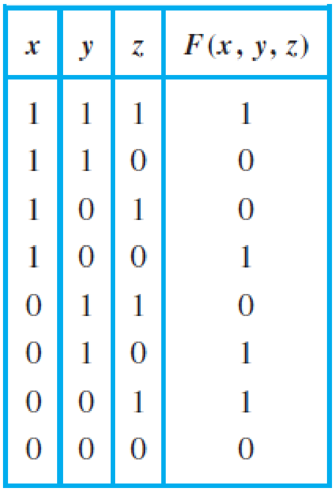
The table displays these values.

x	y	$F(x, y)$
1	1	1
1	0	0
0	1	0
0	0	1

Note that $F(x, y) = xy + x\ y$.

This function is implemented by the circuit shown in the figure.

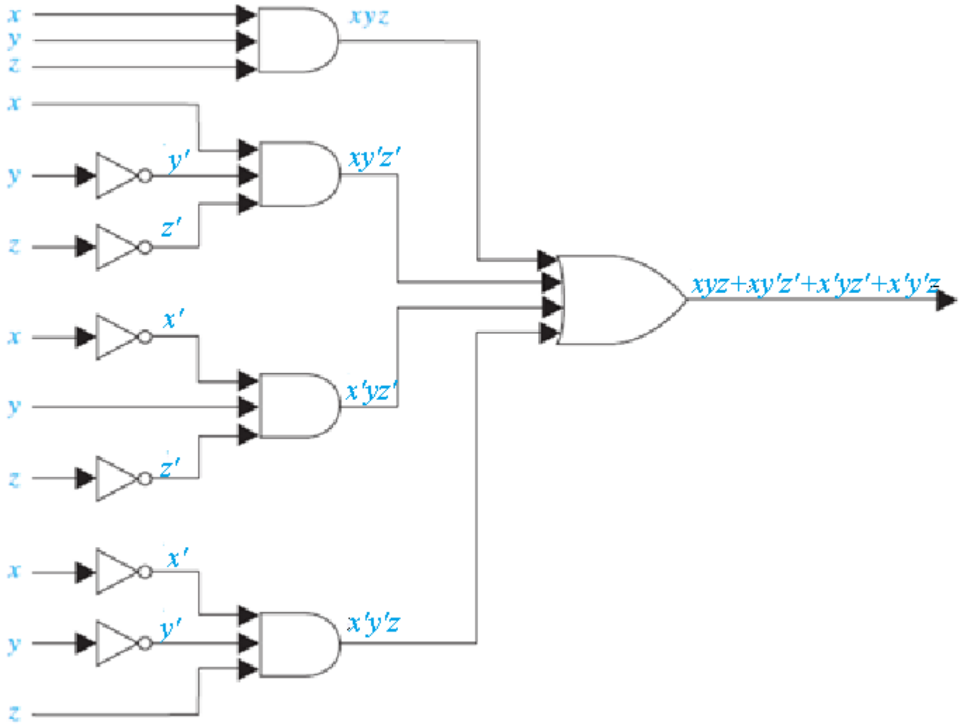

x
y
xy
x
$\overline{x}$
y
$\overline{y}$
$\overline{x}\,\overline{y}$
$xy + \overline{x}\,\overline{y}$



x	y	z	$F(x, y, z)$
1	1	1	1
1	1	0	0
1	0	1	0
1	0	0	1
0	1	1	0
0	1	0	1
0	0	1	1
0	0	0	0

We will now design a circuit for three switches. Let x, y, and z be the Boolean variables that indicate whether each of the three switches is closed. We let $x = 1$ when the first switch is closed, and $x = 0$ when it is open; $y = 1$ when the second switch is closed, and $y = 0$ when it is open; and $z = 1$ when the third switch is closed, and $z = 0$ when it is open. Let $F(x, y, z) = 1$ when the light is on and $F(x, y, z) = 0$ when the light is off. We can arbitrarily specify that the light be on when all three switches are closed, so that $F(1, 1, 1) = 1$. This determines all other values of F. When one switch is opened, the light goes off, so $F(1, 1, 0) = F(1, 0, 1) = F(0, 1, 1) = 0$. When a second switch is opened, the light goes on, so $F(1, 0, 0) = F(0, 1, 0) = F(0, 0, 1) = 1$. Finally, when the third switch is opened, the light goes off again, so $F(0, 0, 0) = 0$. The table shows the values of this function.

The function F can be represented by its sum-of-products expansion as $F(x, y, z) = xyz\ +\ xy'z'\ +\ x'yz'\ +\ x'\,y'z'$. The circuit shown in the following figure implements this function.



♣Minimization of Circuits

Example.

Represent the Boolean function by logic circuit:

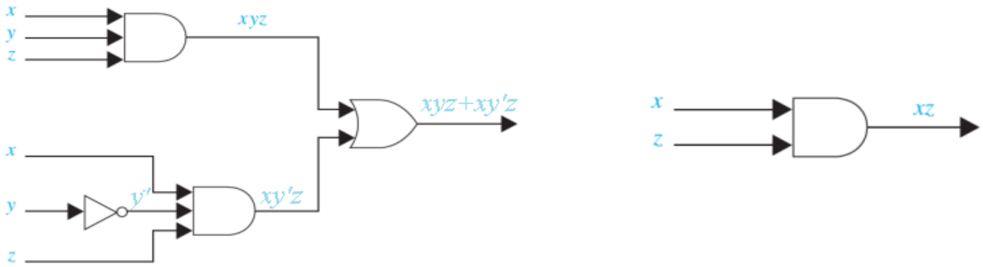
$$F(x, y, z) = xyz + xy'z$$

Solution.

The sum-of-products expansion of this circuit is $xyz + xy'z$. The two products in this expansion differ in exactly one variable, namely, y. They can be combined as

$xyz + xy'z = x(y + y')z = x \cdot 1 \cdot z = xz.$

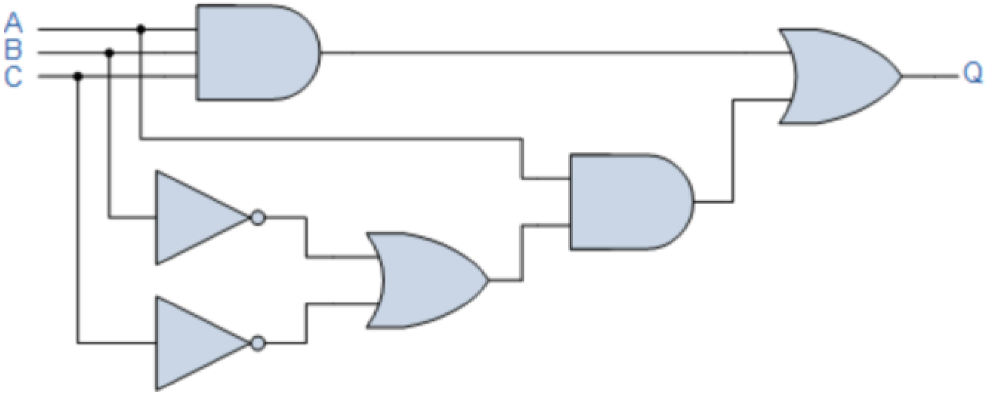
Hence, xz is a Boolean expression with fewer operators that represents the circuit. We show two different implementations of this circuit in the figure. The second circuit uses only one gate, whereas the first circuit uses three gates and an inverter.



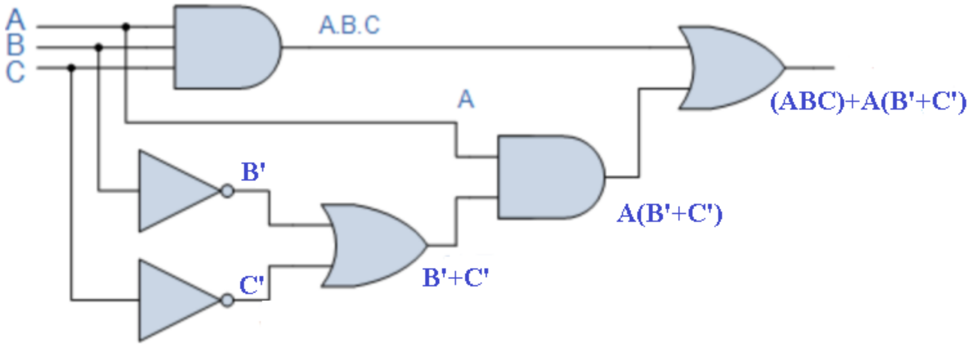
This example shows that combining terms in the sum-of-products expansion of a circuit leads to a simpler expression for the circuit.■

Example.

Find the Boolean algebra expression for the following system.



Solution.



Exercises set (7)

1- Find the values of these expressions

$10'$; $1+1'$; $0'0$; $(1-0)'$.

2- Find the values, if any, of the Boolean variable *x* that satisfy these equations

$x.1=0$, $x+x=0$, $x.1=x$, $x.x'=1$.

3- Use a table to express the values of each of these Boolean functions

$F(x,y,z)=x'y, F(x,y,z)=x+yz$

$F(x,y,z)=xy'+(xyz)'$,

$F(x,y,z)=x(yz+(yz)')$.

4- Find the duals of these Boolean expressions

$x+y, x'y', xyz+(xyz)', xz'+x0+x'$

5- Find a Boolean product of the Boolean variables *x*, *y* and *z* or their complements, that has the value 1 if and only if

(a) $x=y=0,\ z=1$;

(b) $x=0,\ y=1,\ z=0$;

(c) $x=0,\ y=z=1$;

(d) $x=y=z=0$.

6-Find the sum-of-products expressions of these Boolean functions

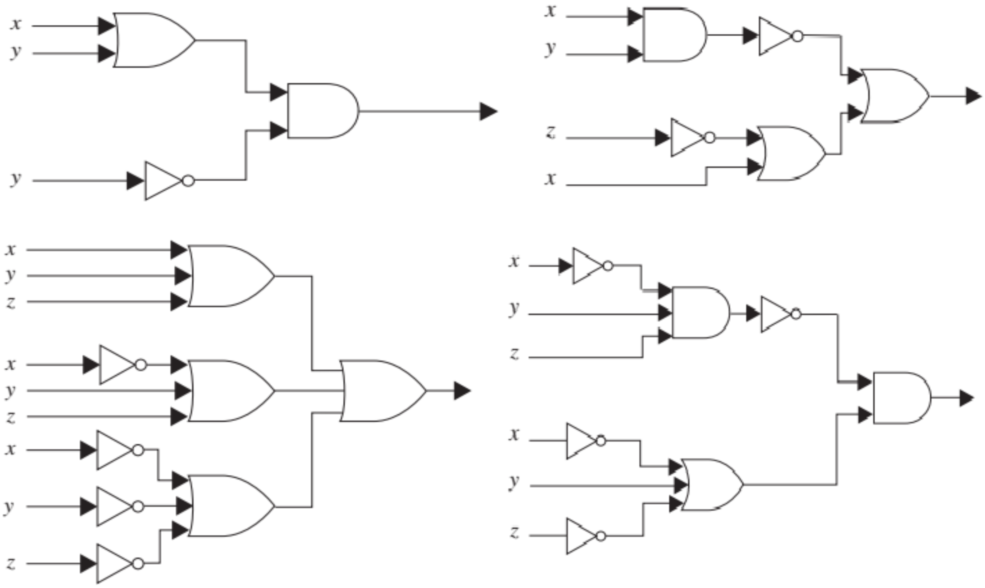
$F(x, y) = x' + y$; $F(x, y) = xy'$; $F(x, y) = 1$;

$F(x, y) = y'$; $F(x, y, z) = x + y + z$;

$F(x, y, z) = (x + z)y$; $F(x, y, z) = x$.

7-Find the products-of-sums expressions of these Boolean functions in Exercise 6.

8-Find the output of the given circuit.



9. Construct circuits to produce these outputs:

a. $x + y$; b. $(x + y)x$; c $xyz + xyz$; d. $(x + z)(y + z)$.

10. Design a circuit that implements majority voting for five individuals.

CHAPTER (VIII)

GRAPH THEORY

Chapter (VIII)

Graph Theory

8.1 Introduction

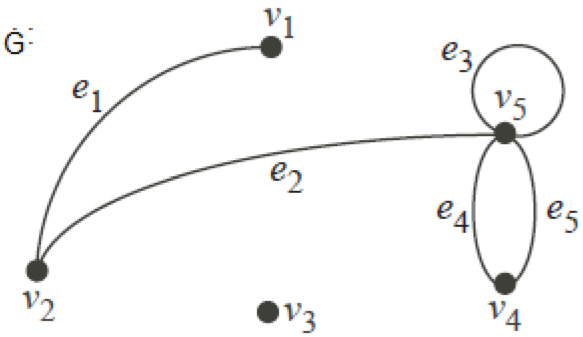
Graphs are discrete structures consisting of vertices and edges that connect these vertices. Problems in almost every conceivable discipline can be solved using graph models. Using graph models, we can determine whether it is possible to walk down all the streets in a city without going down a street twice, and we can find the number of colors needed to color the regions of a map. Graphs can be used to determine whether two computers are connected by a communications link using graph modules of computer networks. Also, graphs can be used to determine whether a circuit can be implemented on a planner circuit board. Graph with weights assigned to their edges can be used to solve problems such as finding the shortest path between two cities in a transportation network.

This chapter will introduce the basic concepts of graph theory and present many different graph models.

8.2 Graphs and Graph Models

Definition.

Conceptually, a ***graph*** is formed by ***vertices*** and ***edges*** connecting the vertices.



Formally. Let V be a non-empty set, E be another set, and f be a mapping such that $f: E \to \{\{x, y\}: x, y \in V\}$. Then the triple $G = (V, E, f)$ is called a ***graph***.

We call that V (or $V(G)$) the set of ***vertices*** of G and E (or $E(G)$) the set of ***edges*** (lines) of G. The graph $G = (V, E, f)$ is finite if each V and E is finite. We consider only the *finite graphs* without explicitly state.

☻ If $v \in f(e)$, then v is an vertex for e.

☻ If $a, b \in V$, then a is ***adjacent*** to b if there exists $e \in E$ such that $f(e) = \{a, b\}$.

☻ Also, $a \in V$ is adjacent to itself if there exists $e \in E$ such that $f(e) = \{a\}$ and e is called a ***loop*** at a.

☻ If $e_{1,}e_2 \in E$ are incident with a common vertex, then we say e_1 and e_2 ***adjacent edges***.

☻If $f(e_1) = f(e_2) = \{a, b\}$, then e_1 and e_2 are called **a *multiple edge*.**

☻ **If** $f(e_1) = f(e_2) = \{v\}$, then e_1 and e_2 are called **a** ***multiple loop*** at v.

☻ A graph G with no loops and no multiple edges is **a** ***simple graph***.

☻ If $G = (V, E, f)$ is a graph and $f(e) = \{a, b\}$, then we write $e = \{a, b\}$ and so we write $G = (V, E)$ instead of $G = (V, E, f)$.

We sometimes consider the following generalizations of graphs: a ***multigraph*** is a pair (V, E) where V is a set and E is a ***multiset*** of unordered pairs from V . In other words, we allow more than one edge between two vertices. A ***pseudograph*** is a pair (V, E) where V is a set and E is a ***multiset*** of unordered multisets of size two from V . A pseudograph allows ***loops***, namely edges of the form $\{a, a\}$ for $a \in V$.

☻ In general, we visualize graphs by using points to represent vertices and line segments, possibly curved, to represent edges.

Definition.

The set of all neighbors of a vertex v of $G = (V, E)$, denoted by $N(v)$, is called the neighborhood of v. If A is a subset of V , we denote by $N(A)$ the set of all vertices in G that are adjacent to at least one vertex in A. So, $N(A) \ = \bigcup_{v \in A} N(v)$.

To keep track of how many edges are incident to a vertex, we make the following definition.

Definition.

Let $G = (V, E)$ be a graph and $x \in V$. The **degree** of x (denoted by $d_G(\ x)$) is the number of edges incident with it, except a loop at x contributes twice to the degree of x.

☻ If $d_G(x) = 0$, then x is said to be **isolated** vertex.

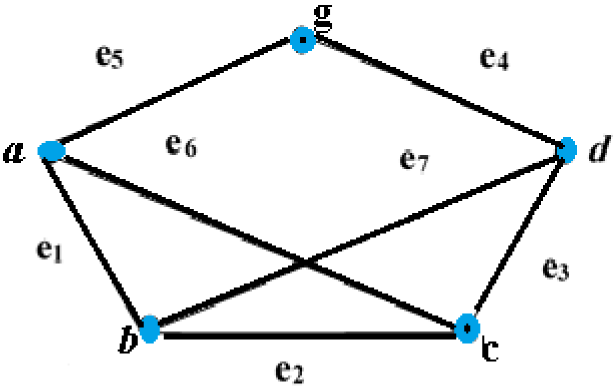
☻ A vertex is **pendant** if and only if it has degree one.

☻ A vertex with odd degree is said to be **odd vertex** and one with even degree is said to be **even vertex**.

☻ The degree sequence of a graph G is the ***sequence*** of degrees of vertices of G in non-increasing order.

Note.

We represent a graph by means of a diagram.



Graph H:

Thus, in the graph H:

☻ The points a and b are adjacent, but a and d are not.

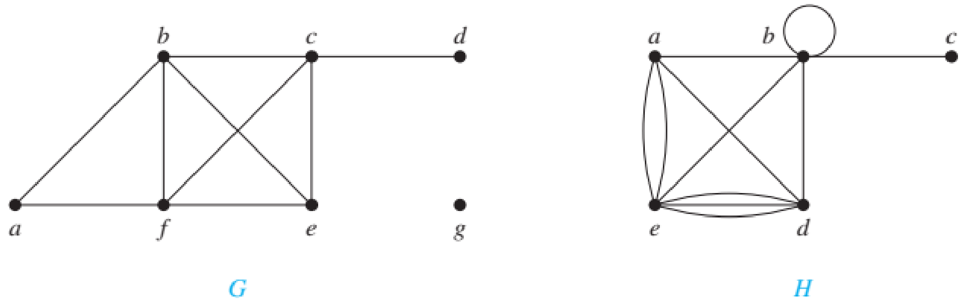
☻ The lines e_2 and e_6 are adjacent but e_6 and e_7 are not.

☻Although the lines e_6 and e_7 are intersect in the diagram but their **intersection** is not a vertex of the graph.

☻ The degree sequence of the graph H is (3,3,3,3,2).

Example.

What are the degrees and what are the neighborhoods of the vertices in the graphs G and H displayed in the given figure?

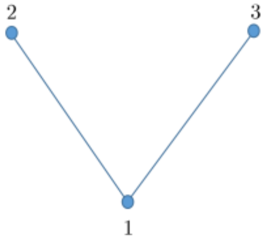


Solution.

In G, $d_G(a) = 2$, $d_G(b) = d_G(c) = d_G(f) = 4$, $d_G(d) = 1$, $d_G(e) = 3$, and $d_G(g) = 0$. The neighborhoods of these vertices are $N(a) = \{b, f\}$, $N(b) = \{a, c, e, f\}$, $N(c) = \{b, d, e, f\}$, $N(d) = \{c\}$, $N(e) = \{b, c, f\}$, $N(f) = \{a, b, c, e\}$, and $N(g) = \phi$.

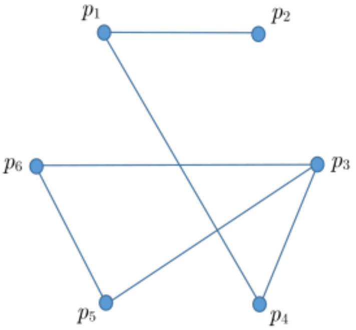
In H, $d_H(a) = 4$, $d_H(b) = d_H(e) = 6$, $d_H(c) = 1$, and $d_H(d) = 5$. The neighborhoods of these vertices are $N(a) = \{b, d, e\}$, $N(b) = \{a, b, c, d, e\}$, $N(c) = \{b\}$, $N(d) = \{a, b, e\}$, and $N(e) = \{a, b, d\}$.■

Example.



Consider the graph $G = (V, E)$, where $V = \{1, 2, 3\}$ and $E = \{\{1, 2\}, \{1, 3\}\}$. Then the given drawing represents this graph.■

Example.



Let $V = \{p_1, p_2, p_3, p_4, p_5, p_6\}$ be a set of six people at a party, and suppose that p_1 shook hands with p_2 and p_4, p_3 shook hands with p_4; p_5 and p_6, and p_5 and p_6 shook hands. Let $G = (V, E)$ be the graph with edge set E consisting of pairs of people who shook hands. Then

$E = \{\{p_1, p_2\}, \{p_1, p_4\}, \{p_3, p_4\}, \{p_3, p_5\}, \{p_3, p_6\}, \{p_5, p_6\}\}$

A drawing of G is given in given figure. ■

Example.

Let $\mathbb{Z}$ denote the set of integers and let

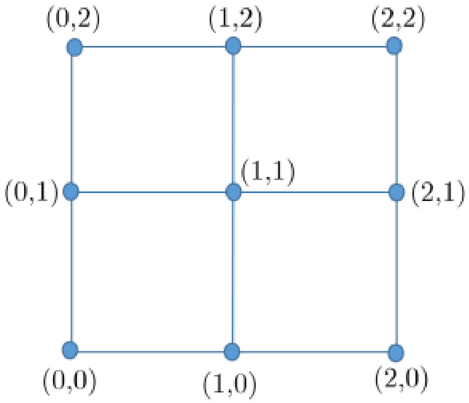
$V = \{(x, y) \in \mathbb{Z} \times \mathbb{Z} : 0 \le x \le 2, 0 \le y \le 2\}$:

Then V is just the set of points in the plane with integer co-ordinates between 0 and 2. Now, suppose $G = (V, E)$

is the graph where E is the set of pairs of vertices of V at distance 1 from each other. In other words, (x, y) and (x', y') are adjacent iff $(x - x')^2 + (y - y')^2 = 1$.

We check that the edge set is

$E \ = \{\{(0,0)(0,1)\}, \{(0,0)(1,0)\}, \{(0,1)(0,2)\},$
$\{(1,0)(2,0)\}, \{(1,0)(1,1)\}, (\{1,1)(1,2)\}, \{(1,1), (2,1)\},$
$\{(0,1), (1,1)\}, \{(0,2)(1,2)\}, \{(2,0)(2,1)\}, \{(2,1), (2,2)\},$
$\{(1,2), (2,2)\}\}$:



This is a cumbersome way to write the edge set of G, as compared to the drawing of G in the given figure, which is much easier to absorb. The graph is called **grid** graph. ■

Example.

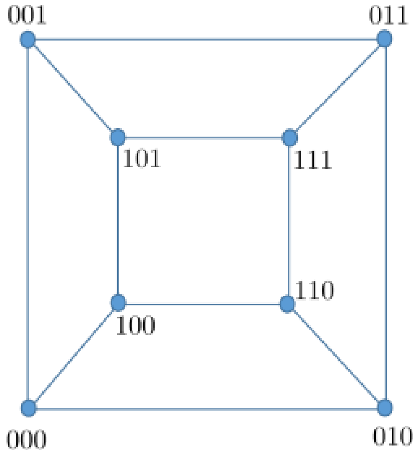
Let V be the set of binary strings of length three, so $V = \{000, 001, 010, 100, 011, 101, 110, 111\}$:

Then let E be the set of pairs of strings which differ in one position. Then

$E = \{\{000, 001\}, \{010, 000\}, \{100, 000\}, \dots, \{111, 101\},$

$\{111, 110\}, \{111, 011\}\}$:

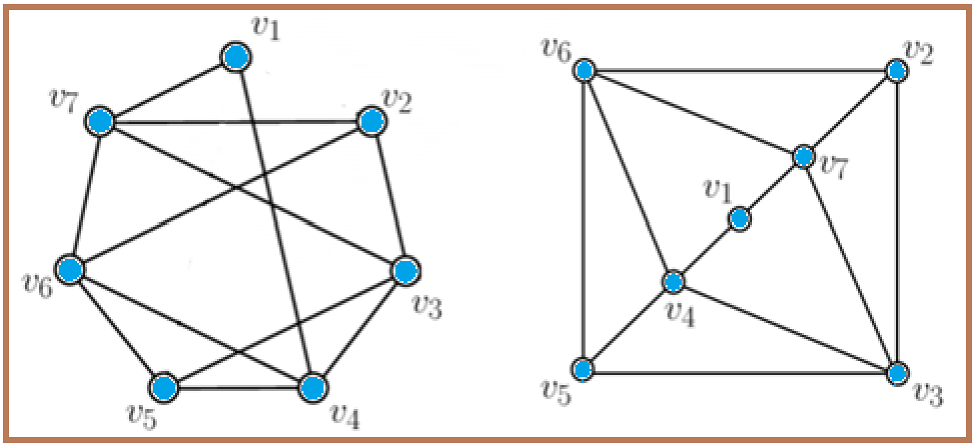
The reader should fill in the rest of the edges as an exercise. Once again, this graph actually has a very nice drawing (which explains why it is sometimes called the **cube** graph).



Example.

Consider the graph $G = (V, E)$, where the vertex set is $V = \{v_1, v_2, v_3, v_4, v_5, v_6, v_7\}$ and the edge set is $E = \{\{v_1, v_4\}, \{v_1, v_7\}, \{v_2, v_3\}, \{v_2, v_6\}, \{v_2, v_7\}, \{v_3, v_4\}, \{v_3, v_5\}, \{v_3, v_7\}, \{v_4, v_5\}, \{v_4, v_6\}, \{v_5, v_6\}, \{v_5, v_7\}\}$:

In the following figure, two drawings of G are shown (the reader should verify that they are both drawings of G)



Example.

Let $G = (V, E)$ be a graph, where $V = \{a, b, c, d, g\}$, $E = \{e_1, e_2, e_3, e_4, e_5, e_6\} = \{\{a\}, \{a, b\}, \{a, c\}, \{a, c\}, \{b, c\},$ $\{c, d\}\}$

1. Represent the graph G;

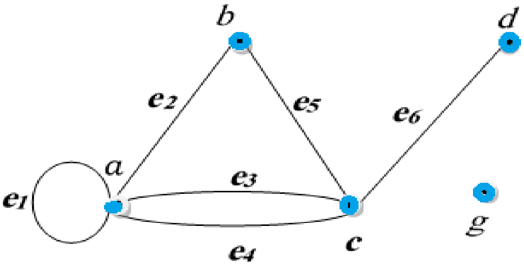
2. Find the degree of each vertex and isolated vertices;

3. Find multiple edges and loops;

4. Is G a simple graph? Why?

Solution.

1.



Graph G:

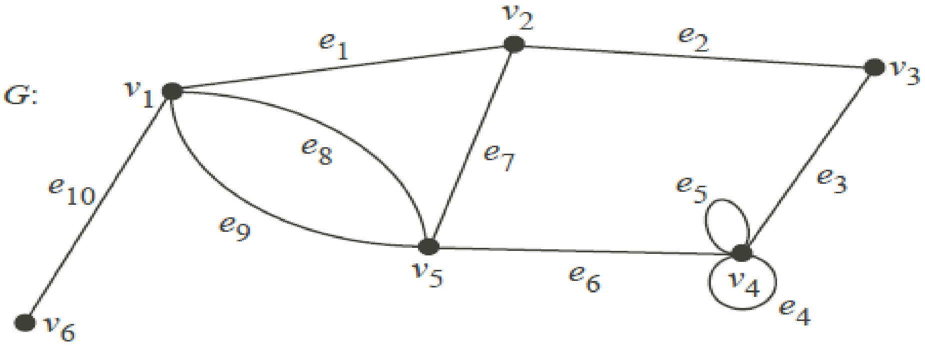
2. $d_G(a) = 5, d_G(b) = 2, d_G(c) = 4, d_G(d) = 1,$ $d_G(g) = 0$. Therefore the degree sequence is $(5, 4, 2, 1, 0)$. Since $d_G(g) = 0$ then g is the only isolated vertex.

3. Since $e_3 = e_4 = \{a, c\}$, e_3 and e_4 are multiple edges and hence G is a multiple graph. Also, since $e_1 = \{a\}$, then e_1 is a loop.

4. G is not a simple graph. It is a pseudograph as it contains multiple edges and a loop. ■

Example.

If $G = (V, E, f)$ is the graph given by the following diagram



Find V, E, f.

Solution.

It is clear that $V = \{v_1, v_2, v_3, v_4, v_5, v_6\}$. and $E = \{e_1, e_2, e_3, e_4, e_5, e_6, e_7, e_8, e_9, e_{10}\}$.

The following table represents the function *f*:

E	e_1	e_2	e_3	e_4	e_5
$f(e)$	$\{v_1, v_2\}$	$\{v_2, v_3\}$	$\{v_3, v_4\}$	$\{v_4\}$	$\{v_4\}$

E	e_6	e_7	e_8	e_9	e_{10}
$f(e)$	$\{v_4, v_5\}$	$\{v_5, v_2\}$	$\{v_1, v_5\}$	$\{v_1, v_5\}$	$\{v_1, v_6\}$

.■

Definition.

We write $\delta(G) = \min\{d_G(v): v \in V\}$ and $\Delta(G) = \max\{d_G(v): v \in V\}$ for the ***minimum degree*** and ***maximum degree*** of G, respectively.

Note.

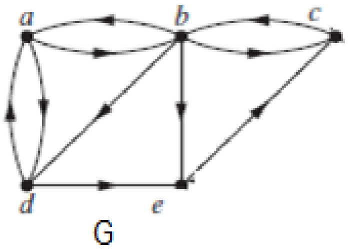
The graphs we have introduced are **undirected graphs**. Their edges are also said to be undirected. To construct a graph model, we may find it is necessary to assign direction to the edges of a graph.

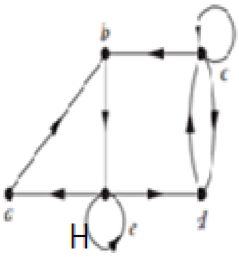
Definition.

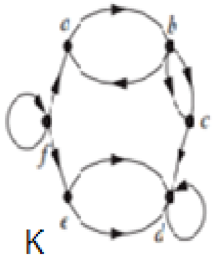
A directed graph (or digraph) $G = (V, E, f)$ consists of a non-empty set of vertices V and set of directed edges (or arcs) with the map $f: E \to \{(x, y): x, y \in V\}, i.e.$, each directed edge is associated with an ordered pair of vertices. The directed edge associated with the ordered pair (u, v) is said to start at u and end at v. If $f(e_1) = f(e_2)$ in digraph, then e_1 and e_2 are multiple edges. If a digraph G contains no multiple edges or graph loops, then it a *directed simple graph***.**

Example.

G is a simple directed graph while H and K are not.



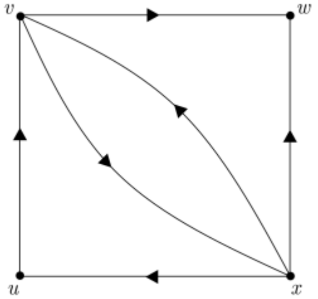




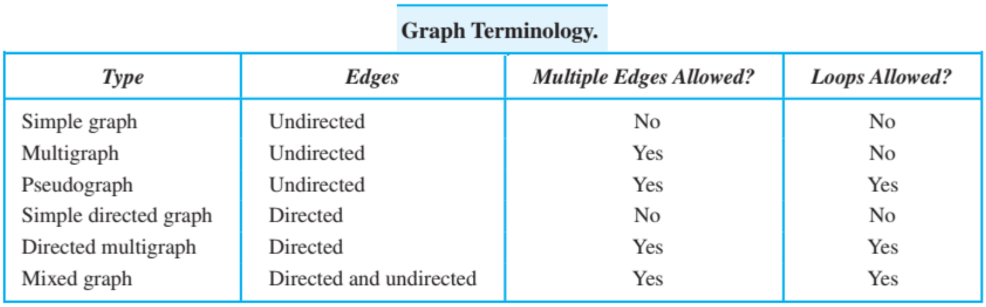
Note:

(a) If $e = (u, v)$ is an edge of a digraph G, then u is the *initial* vertex and v is the *terminal* vertex for the edge e.

(b) In a digraph G, let $N^+(v)$ and $N^-(v)$ denote the sets of vertices adjacent from v and to v, respectively. These are the *out-neighborhood* of v and the *in-neighborhood* of v respectively. Thus $N^+(v) = \{u: (v, u) \in E\}$ and $N^-(v) = \{u: (u, v) \in E\}$. For example, in the digraph drawn below, $N^+(x) = \{u, v, w\}$ and $N^-(x) = \{v\}$.



(c) A graph with both directed and undirected edge is called a **mixed** graph.



Graph Terminology.

Type	*Edges*	*Multiple Edges Allowed?*	*Loops Allowed?*
Simple graph	Undirected	No	No
Multigraph	Undirected	Yes	No
Pseudograph	Undirected	Yes	Yes
Simple directed graph	Directed	No	No
Directed multigraph	Directed	Yes	Yes
Mixed graph	Directed and undirected	Yes	Yes

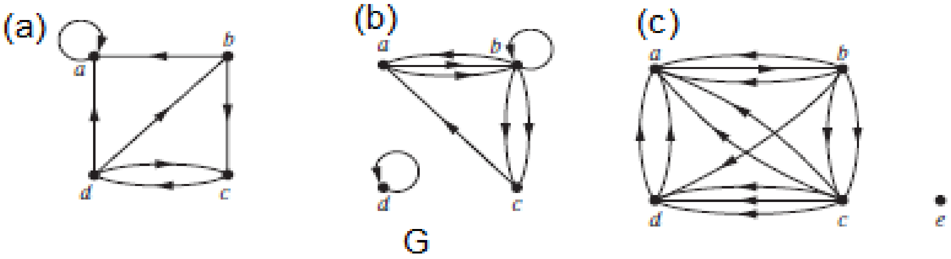
Definition.

In a graph with directed edge the **in-degree** of a vertex v, denoted by (or $d_G^-(v)$) is the number of edges with v as their terminal vertex. The **out-degree** of a vertex v denoted by (or $d_G^+(v)$)) is the number of edges with v as their initial vertex. A loop at v contributes one to the in-degree and one to the out-degree of v. In other words,

$d_G^-(v) = |N^-(v)|$ and $d_G^+(v) = |N^+(v)|$.

Example.

Find the in-degree and out-degree of each vertex in the digraph G Shown in the following diagram.



Solution.

The following tables gives the out-degree and in-degree of each vertex in Graphs G-(a), G-(b) and G-(c), respectively.

G-(a): G-(b):

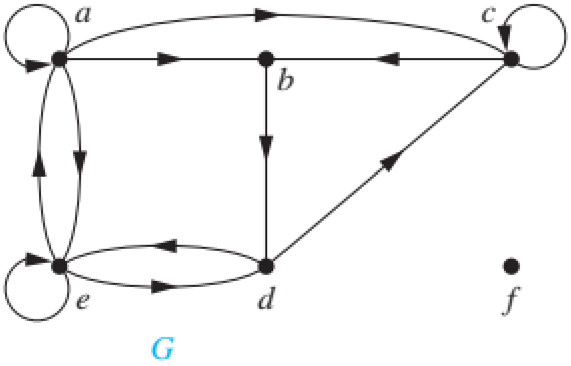
v	a	b	c	d		a	b	c	d
$d_G^-(v)$	3	1	2	1		2	3	2	1
$d_G^+(v)$	1	2	1	3		2	4	1	1

G-(c):

v	a	b	c	d	e
$d_G^-(v)$	6	1	2	4	0
$d_G^+(v)$	1	4	5	2	0

Example.

Find the in-degree and out-degree of each vertex in the graph G with directed edges shown in the given Figure.



Solution.

The in-degrees in G are $d_G^-(a) = 2$, $d_G^-(b) = 2$, $d_G^-(c) = 3$, $d_G^-(d) = 2$, $d_G^-(e) = 3$, and $d_G^-(f) = 0$. The out-degrees are $d_G^+(a) = 4$, $d_G^+(b) = 1$, $d_G^+(c) = 2$, $d_G^+(d) = 2$, $d_G^+(e) = 3$, and $d_G^+(f) = 0$. ■

Because each edge has an initial vertex and a terminal vertex, the sum of the in-degrees and the sum of the out-degrees of all vertices in a graph with directed edges are the same. Both of these sums are the number of edges in the graph. This result is stated as the following theorem.

The following theorem is called **Handshaking Theorem.** It describes the relation between the number of edges of a graph and the degrees of its vertices.

Theorem.

Let $G = (V, E)$ be a graph such that $V = \{x_1, \dots, x_n\}$.

Then

(a) $\sum_{i=1}^{n} d_G(x_i) = 2|E|$;

(b) The number of odd vertices in G is even;

(c) In a digraph $G, \sum_{i=1}^{n} d_G^-(x_i) = \sum_{i=1}^{n} d_G^+(x_i) = |E|$.

Proof.

(a) We compute the number of times that edges of G are incident with its vertices by two different ways.

First, each edge is incident with vertices twice, i. e., the desired number is $2|E|$. In other words, each vertex is incident with edges $(d_G(x))$ once. Therefore, The desired number is $\sum_{i=1}^{n} d_G(x_i)$. Thus $\sum_{i=1}^{n} d_G(x_i) = 2|E|$.

(b) Let V_1 and V_2 de the set of vertices of even degree and set of vertices of odd degree, respectively, in G. Then $V = V_1 \cup V_2$ and $V_1 \cap V_2 = \emptyset$. Therefore

$\sum_{i=1}^{n} d_G(x_i) = \sum_{x \in V_1} d_G(x) + \sum_{x \in V_2} d_G(x) = 2|E|$.

Since both $2|E|$and $\sum_{x \in V_1} d_G(x)$ are even, then $\sum_{x \in V_2} d_G(x)$ is even. Since all terms in this sum is odd, then there must be an even number of such terms. Thus there is an even number of vertices of odd degree.

(c) Since each edge has an initial vertex and a terminal vertex, then the sum of the in-degrees and the sum of the out-degrees of all vertices in a graph with directed edges are the same. Both sums are the number of edges $|E|$ in the graph.■

Example.

Consider the grid graph. The degree sequence of this graph is $(4,\ 3,\ 3,\ 3,\ 3,\ 2,\ 2,\ 2,\ 2)$. Therefore by the handshaking theorem, the number of edges in the grid graph is: $\frac{1}{2}(4 + 3 + 3 + 3 + 3 + 2 + 2 + 2 + 2) = 12$.

A manual count of the edges in the grid graph confirms this. The reader should check how many edges the n by n

grid graph has (the vertex set is $V = \{(x, y) \in \mathbb{Z} \times \mathbb{Z} : 0 \leq x < n, 0 \leq y < n\}$ and the edge set is the set of pairs of vertices at distance 1 from each other.) ■

Example.

1. How many edges are there in a graph with 10 vertices each of degree six?

2. Is there a graph the sequence of degrees of its vertices is $(5, 4, 3,\ 3, 2)$?

Solution.

1. Since the degrees of the vertices is $6 \times 10 = 60 = 2|E|$, then $|E| = 60/2 = 30$.

2. Since $5 + 4 + 3 + 3 + 2 = 17$ is an odd number, by the handshaking theorem, there is no graph with these vertices. Or, since the number of the odd vertices is 3, then there is no graph with these vertices by the same theorem.■

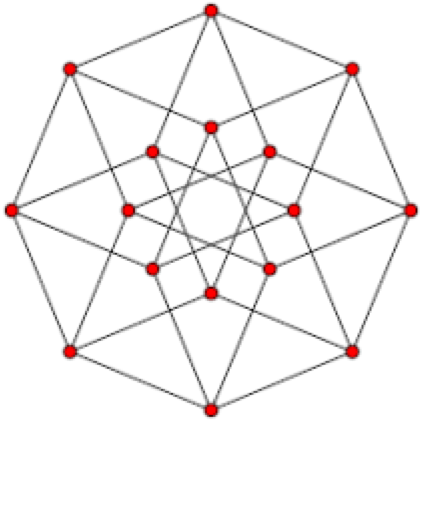
Example.

The *n*-cube, denoted Q_n, is the graph whose vertex set is the set of binary strings of length n, and whose edge set consists of all pairs of strings differing in one position.

The cube graph Q_3 in introduced in this section is the 3-cube. Let us see how many edges Q_n has as a formula in n. Since there are 2^n binary strings of length n, there are 2^n vertices in Q_n. Now each vertex v is adjacent to n other vertices - namely flip one position in the string v to get each string adjacent to v, and there are n possible positions in which to do a flip. So every vertex of the n-cube has degree n, and so the number of edges in Q_n is

$$\frac{1}{2}\sum_{n \in V} d_{Q_n}(v) = \frac{1}{2} \cdot 2^n \cdot n = n2^{n-1}$$

A manual count of the edges confirms this for the 4-cube Q_4 which is drawn below:



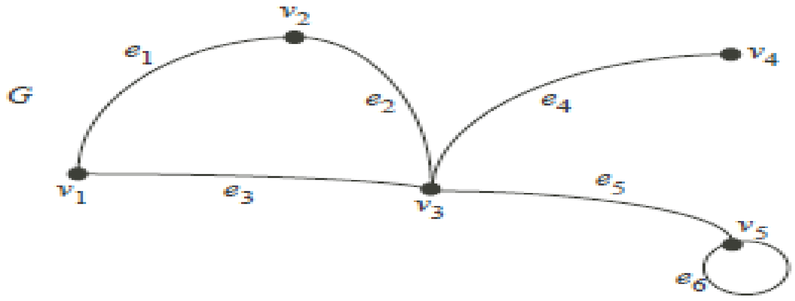
8.3 Subgraph

Definition.

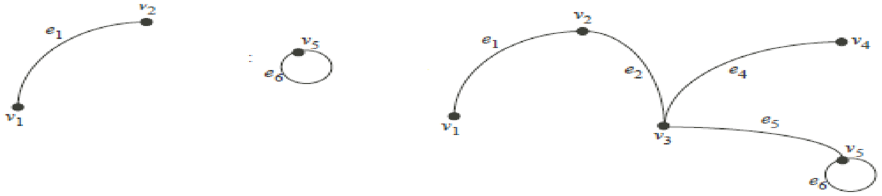
A subgraph of a graph $G = (V, E)$ is a graph $H = (W, F)$, where $W \subseteq V$ and $F \subseteq E$. The subgraph H of the graph G is **spanning** the graph G if $W = V$. If $\{x_1, \dots, x_n\} \subseteq V$ in the graph $G = (V, E)$, we obtain the subgraph $G - \{x_1, \dots, x_n\}$ by deleting the vertices $x_1, \dots, x_n$ and all fallen edges. If $\{e_1, \dots, e_n\} \subseteq E$ in the graph $G = (V, E)$, then we get the subgraph $G - \{e_1, \dots, e_k\}$ by deleting the edges $e_1, \dots, e_k$ (without deleting the vertices).

Example.

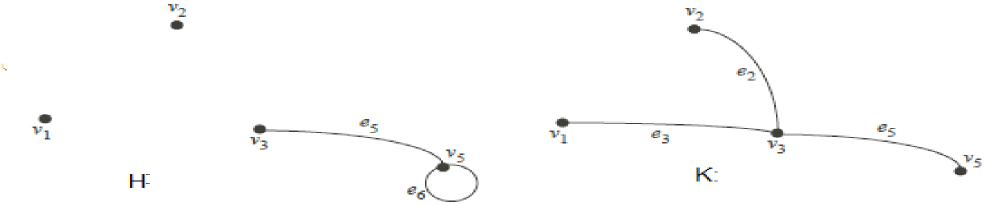
Let G be the following graph.



The following three graphs all subgraphs of G:

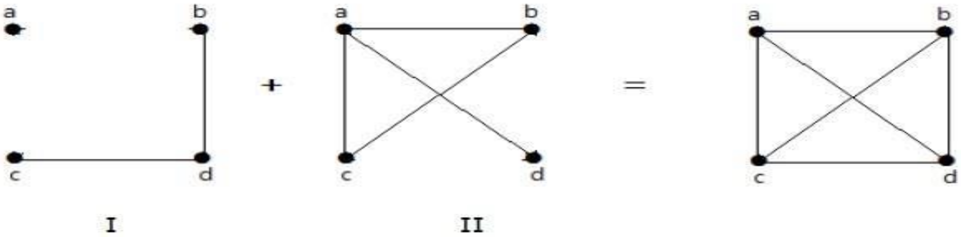


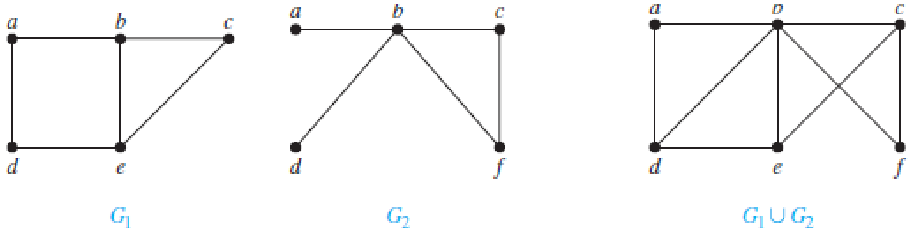
The following is the subgraphs $H = G - \{e_1, e_2, e_3, v_4\}$ and $K = G - \{e_1, e_6, v_4\}$.



Definition.

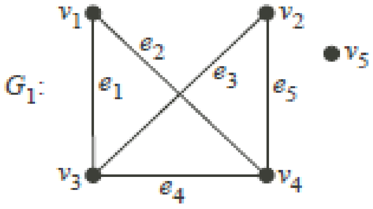
The union of two simple graphs $G_1 = (V_1, E_1)$ and $G_2 = (V_2, E_2)$ is the simple graph with vertex set $V_1 \cup V_2$ and edge set $E_1 \cup E_2$. The union of G_1 and G_2 is denoted by $G_1 \cup G_2$.

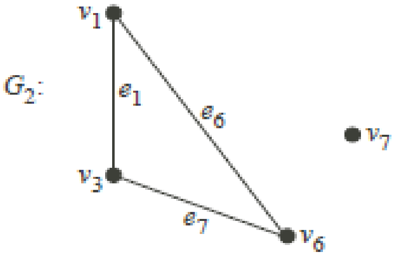




Example.

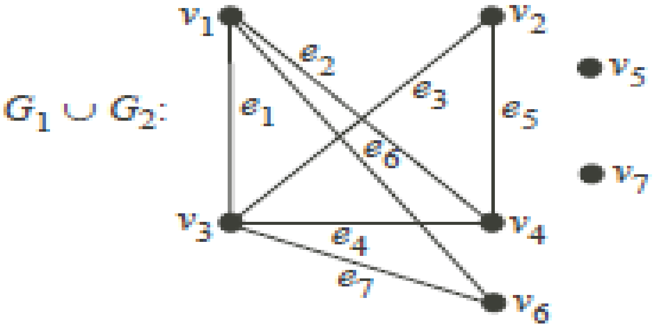
Find the union of the following graphs.





Solution.

The vertex set of the union $G_1 \cup G_2$ is the union of the two vertex sets, namely $\{v_1, v_2, v_3, v_4, v_5, v_6, , v_7\}$. The edge set of $G_1 \cup G_2$ is the union of the two edge sets, namely $\{e_1, e_2, e_3, e_4, e_5, e_6, e_7\}$. The union is displayed in following figure.

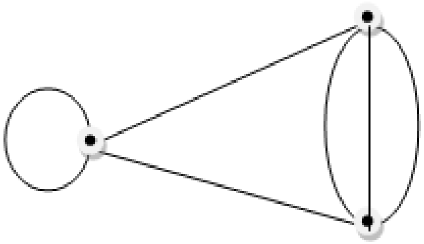


8.4 Special Graphs

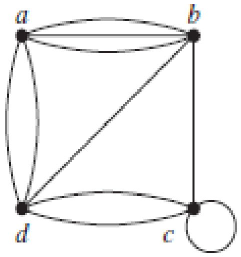
Definition.

Let $G = (V, E)$ be a graph and $r \geq 0$ be an integer. The graph G is said to be **r-regular** graph if $d_G(x) = r$ for each $x \in V$. For instance, the graph Q_3 is 3-regular (all the degrees are 3). Sometimes, 3-regular graphs are also referred to as *cubic* graphs.

Example.



4 - regular graph



5 - regular graph

Theorem.

If $G = (V, E)$ is r-regular graph with $|V| = n$, $|E| = \frac{nr}{2}$.

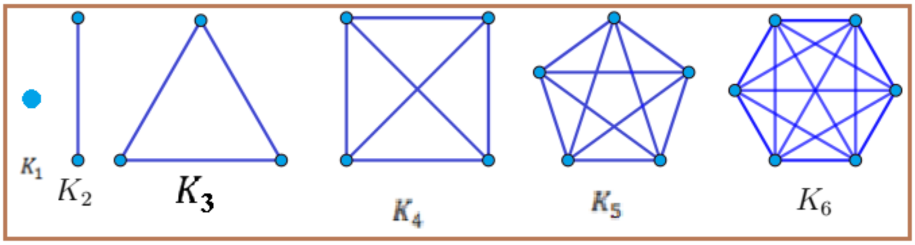
Proof.

Since $\sum_{x \in V} d_G(x) = 2|E|$, Then $\sum_{x \in V} r = 2|\mathrm{E}|$.

Therefore $nr = 2|E|$ or $|E| = \frac{nr}{2}$.■

Definition.

The complete graph with n vertices, denoted by K_n is the simple graph that contains exactly one edge between every pair of distinct vertices.



Theorem.

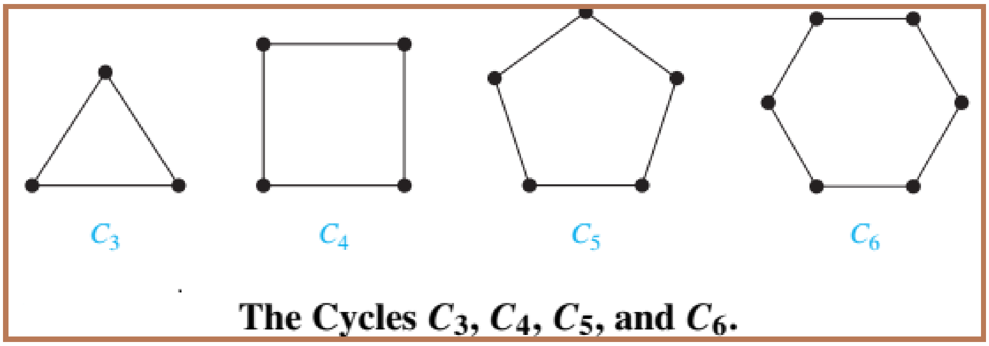
If $K_n = (V, E)$, then $|E| = \frac{n(n-1)}{2}$.

Proof.

K_n is $(n-1)$–regular graph. Therefore $|E| = \frac{n(n-1)}{2}$. ■

Definition. (Cycles).

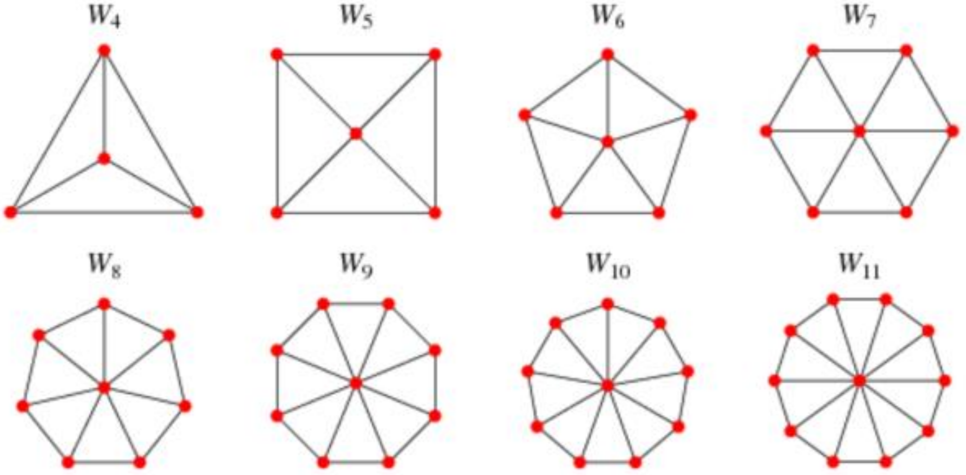
The cycle $C_n, n \geq 3$ consists of n vertices $v_1, \dots, v_n$ and edges $\{v_1, v_2\}, \{v_2, v_3,\}, \dots, \{v_{n-1}, v_n\}$ and $\{v_n, v_1\}$.



The Cycles C_3, C_4, C_5, and C_6.

Definition. Wheels.

We obtain the wheel W_n when we add an additional vertex to the cycle C_n , for $n \geq 3$ and connect this new vertex to each of the n vertices in C_n by new edges.
In the following figure, the wheel graphs W_n with n vertices are shown for $4 \leq n \leq 11$.



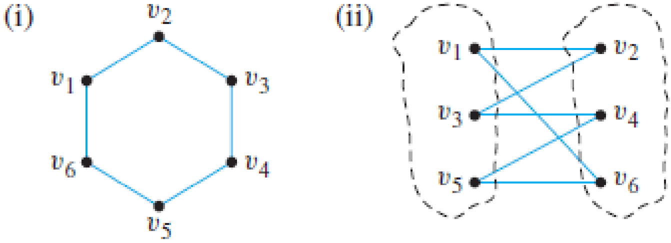
Definition.

A simple graph $G = (V, E)$ is said to be **bipartite graph** if its vertices can be partitioned into two disjoint sets V_1 and V_2 such that every edge in the graph connects a vertex in V_1 and a vertex in V_2 (so that no edge in G connects either two vertices in V_1 or two vertices in V_2). When this condition holds, we call the pair (V_1 , V_2) a *bipartition* of the vertex set V of G.

Note that partition means that $V_1 \neq \phi, V_2 \neq \phi, V_1 \cap V_2 = \phi$ and $V = V_1 \cup V_2$. In this case we use the symbol $(V_1 \cup V_2, E)$ instead of (V, E).

Example.

The graph G in (i) can be redrawn as shown in (ii). From the drawing in (ii), you can see that G is bipartite with mutually disjoint vertex sets $\{v_1, v_3, v_5\}$ and $\{v_2, v_4, v_6\}$.

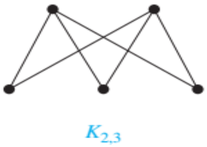


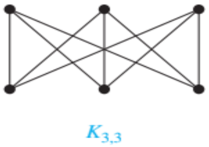
Definition.

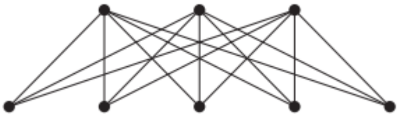
let $G = (V_1 \cup V_2, E)$ be a bipartite graph. G is said to be **complete bipartite** graph if every vertex in V_1 is adjacent to every vertex in V_2. If $|V_1| = m$ and $|V_2| = n$, then this graph is denoted by $K_{m,n}$.

Example.

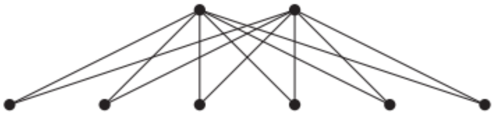
The following graphs are complete bipartite graphs.







$K_{3,5}$



$K_{2,6}$

Theorem.

If $K_{m,n} = (V_1 \cup V_2, E)$ such that $|V_1| = m$ and $|V_2| = n$, then $|E| = mn$.

Proof.

Since $\sum_{x\in v_1} d(x) + \sum_{x\in v_2} d(x) = 2|E|$, then $\sum_{x\in v_1} n + \sum_{x\in v_2} m = 2|E|$. Thus $mn + nm = 2|E|$.

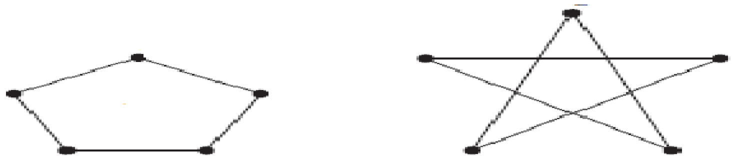
Therefore, $|E| = mn.$ ■

Definition.

Let $G = (V, E)$ be a simple graph. The complement of the graph G is defined to be the graph $\bar{G} = (V, \bar{E})$ where for every $x, y \in V$ and $x \neq y$ we have $\{x, y\} \in \bar{E}$ if and only if $\{x, y\} \notin E$.

Example.

The following diagram is the graph and its complement.



Remark.

If G is r-regular simple graph with n vertices, then $\bar{G}$ is $(n-r-1)-$ regular simple graph.

Exercise.

Give an example of a r- regular simple graph with 6 vertices, where $0 \leq r \leq 5$.

8.5 Representation of Graphs

We will represent graphs using matrices.

The Adjacency Matrix

Definition.

Let $V=\{x_1,x_2,\dots,x_n\}$ and $G=(V,E)$ be a simple graph

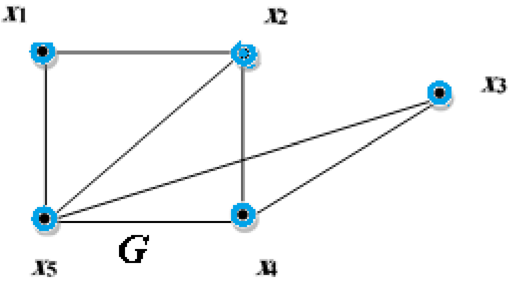
The **adjacency matrix** of the graphs G is the zero - one

matrix $A=[a_{ij}]$, where $a_{ij}=\begin{cases}1, & \{x_i,x_j\}\in E\\ 0, & \{x_i,x_j\}\notin E\end{cases}$

Example.

The adjacency matrix for the given graphs G is:

$$A=\begin{bmatrix}0&1&0&0&1\\1&0&0&1&1\\0&0&0&1&1\\0&1&1&0&1\\1&1&1&1&0\end{bmatrix}$$



Example.

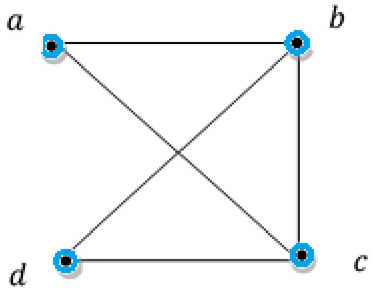
Draw the graph with the adjacency matrix:

$$\begin{bmatrix}0&1&1&0\\1&0&1&1\\1&1&0&1\\0&1&1&0\end{bmatrix}$$

with respect to the ordering of vertices a,b,c,d.

Solution.

The graph with this adjacency matrix is



Remark.

1. The adjacency matrix depends on the ordering of vertices so there exists n ! adjacency matrices for a simple graph with n vertices.

2. The adjacency matrix A of a simple graph G is symmetric, i.e., $A = A^T$, where A^T is the transpose of A**.**

3. Since the simple graph contains no loops, then $a_{ii} = 0$ for every $i \in \{1, \dots, n\}$, i.e., The diameter elements in the adjacency matrix are zeros.

4. We can consider the elements a_{ij} belong to the Boolean algebra $B_2 = \{0,1\}$.

Definition.

Let $G = (V, E)$ be a simple digraph (directed graph), where $= \{x_1, \dots, x_n\}$. Then the **adjacency matrix** for the graph G is the matrix $A = [a_{ij}]$, where

$$a_{ij} = \begin{cases} 1, & (x_i, x_j) \in E \\ 0, & (x_i, x_j) \notin E \end{cases}$$

In this case, A may not be symmetric as it is possible that $(x_i, x_j) \in E$ but $(x_i, x_j) \notin E$.

Definition.

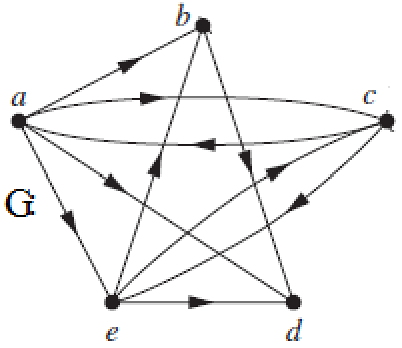
We can define the adjacency matrix of the multi-graph as every loop $\{x_i\}$ participates by one in a_{ii} and every edge $\{x_i, x_j\}, i \neq j$, also , participates by one in a_{ij} . Therefore the elements a_{ij} is not elements of $B_2 = \{0,1\}$.

Definition.

Adjacency matrices can also be used to represent undirected graphs with loops and with multiple edges. A loop at the vertex v_i is represented by a 1 at the $(i, i)^{\text{th}}$ position of the adjacency matrix. When multiple edges connecting the same pair of vertices v_i and v_j, or multiple loops at the same vertex, are present, the adjacency matrix is no longer a zero-one matrix, because the $(i, i)^{\text{th}}$ entry of this matrix equals the number of edges that are associated to $\{v_i, v_j\}$. All undirected graphs, including multigraphs and pseudographs, have symmetric adjacency matrices..

Example.

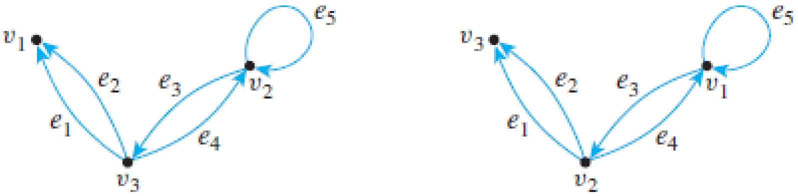
Here the simple digraph G and its adjacency matrix A.



$$A = \begin{bmatrix} 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 \end{bmatrix}$$

Example.

The two directed digraphs shown below differ only in the ordering of their vertices. Find their adjacency matrices.



(a)

(b)

Solution.

Since both graphs have three vertices, both adjacency matrices are 3 × 3 matrices. For (a), all entries in the first row are 0 since there are no arrows from v_1 to any other vertex. For (b), the first two entries in the first row are 1 and the third entry is 0 since from v_1 there are single arrows to v_1 and to v_2 and no arrows to v_3. Continuing the analysis in this way, you obtain the following two adjacency matrices:

$$\begin{array}{c} \\ v_1 \\ v_2 \\ v_3 \end{array}\begin{array}{c} \begin{array}{ccc} v_1 & v_2 & v_3 \end{array} \\ \left[\begin{array}{ccc} 0 & 0 & 0 \\ 0 & 1 & 1 \\ 2 & 1 & 0 \end{array}\right] \end{array} \qquad \begin{array}{c} \\ v_1 \\ v_2 \\ v_3 \end{array}\begin{array}{c} \begin{array}{ccc} v_1 & v_2 & v_3 \end{array} \\ \left[\begin{array}{ccc} 1 & 1 & 0 \\ 1 & 0 & 2 \\ 0 & 0 & 0 \end{array}\right] \end{array}$$

(a) (b)

♣If you are given a square matrix with nonnegative integer entries, you can construct a directed graph with that matrix as its adjacency matrix. However, the matrix does not tell you how to label the edges, so the directed graph is not uniquely determined.

Example.

Draw a directed graph that has **A** as its adjacency matrix.

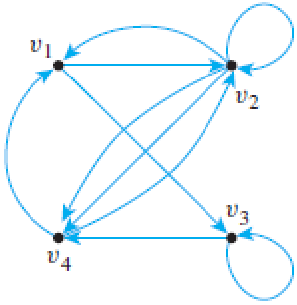
$$\mathbf{A} = \begin{bmatrix} 0 & 1 & 1 & 0 \\ 1 & 1 & 0 & 2 \\ 0 & 0 & 1 & 1 \\ 2 & 1 & 0 & 0 \end{bmatrix}$$

Solution.

Let G be the graph corresponding to **A**, and v_1, v_2, v_3, v_4 be the vertices of G. Label **A** across the top and down the left side with these vertex names, as shown below.

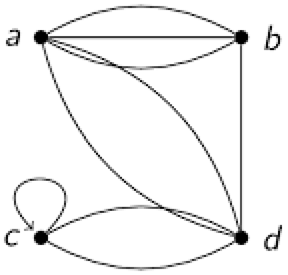
$$\mathbf{A} = \begin{array}{c} \\ v_1 \\ v_2 \\ v_3 \\ v_4 \end{array} \begin{array}{c} \begin{array}{cccc} v_1 & v_2 & v_3 & v_4 \end{array} \\ \begin{bmatrix} 0 & 1 & 1 & 0 \\ 1 & 1 & 0 & 2 \\ 0 & 0 & 1 & 1 \\ 2 & 1 & 0 & 0 \end{bmatrix} \end{array}$$

Then, for instance, the 2 in the fourth row and the first column means that there are two arrows from v_4 to v_1. The 0 in the first row and the fourth column means that there is no arrow from v_1 to v_4. A corresponding directed graph is shown on the next page (without edge labels because the matrix does not determine those). ■



Example.

The adjacency matrix of the given multi-graph is the shown matrix.

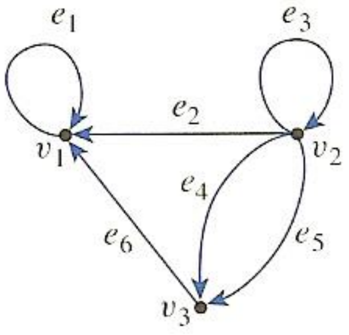


$$\begin{array}{c} \\ a \\ b \\ c \\ d \end{array}\begin{array}{c} \begin{array}{cccc} a & b & c & d \end{array} \\ \begin{pmatrix} 0 & 3 & 0 & 2 \\ 3 & 0 & 0 & 1 \\ 0 & 0 & 1 & 2 \\ 2 & 1 & 2 & 0 \end{pmatrix} \end{array}$$

We used the ordering of vertices $a, b, c\ , d$.■

Example.

The adjacency matrix A for the given multi-digraph G is as follows:



Directed Graph G

$$\mathbf{A} = \begin{array}{c|ccc} & v_1 & v_2 & v_3 \\ \hline v_1 & 1 & 0 & 0 \\ v_2 & 1 & 1 & 2 \\ v_3 & 1 & 0 & 0 \end{array}$$

Adjacency Matrix

Incidence matrices

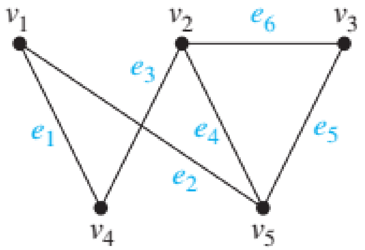
Definition.

Another common way to represent graphs is to use incidence matrices. Let $G = (V, E)$ be an undirected graph. Suppose that $\boldsymbol{v_1}, \boldsymbol{v_2}, \ldots, \boldsymbol{v_n}$ are the vertices and $\boldsymbol{e_1}, \boldsymbol{e_2}, \ldots, \boldsymbol{e_m}$ are the edges of G. Then the incidence matrix with respect to this ordering of V and E is the $n \times m$ matrix $M = [m_{ij}]$, where

$$m_{ij} = \begin{cases} 1 & \text{when edge } \mathrm{e}_j \text{ is incident with } \boldsymbol{v_i}, \\ 0 & \text{otherwise.} \end{cases}$$

Example.

Represent the graph shown in the given figure with an incidence matrix.



Solution.

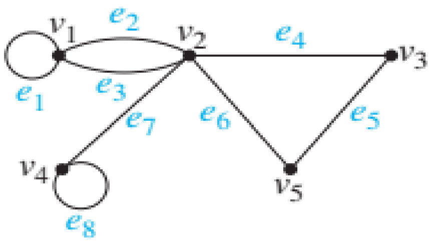
The incidence matrix is:

$$\begin{array}{c} \\ v_1 \\ v_2 \\ v_3 \\ v_4 \\ v_5 \end{array} \begin{array}{c} \begin{array}{cccccc} e_1 & e_2 & e_3 & e_4 & e_5 & e_6 \end{array} \\ \begin{bmatrix} 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 1 & 0 \end{bmatrix} \end{array} .$$

Incidence matrices can also be used to represent multiple edges and loops. Multiple edges are represented in the incidence matrix using columns with identical entries, because these edges are incident with the same pair of vertices. Loops are represented using a column with exactly one entry equal to 1, corresponding to the vertex that is incident with this loop.

Example.

Represent the pseudograph shown in the given figure using an incidence matrix.



Solution.

The incidence matrix for this graph is

$$\begin{array}{c} \\ v_1 \\ v_2 \\ v_3 \\ v_4 \\ v_5 \end{array}\begin{array}{c} \begin{array}{cccccccc} e_1 & e_2 & e_3 & e_4 & e_5 & e_6 & e_7 & e_8 \end{array} \\ \left[\begin{array}{cccccccc} 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 \end{array}\right] \end{array}.$$

8.6 Isomorphism of Graphs

Definition.

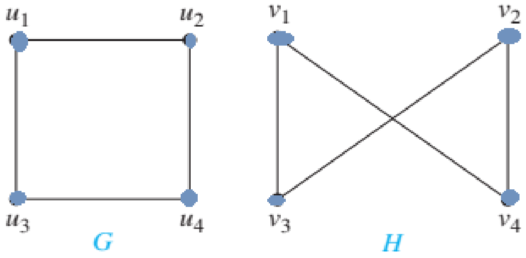
Let $G = (V(G), E(G))$ and $H = (V(H), E(H))$ be two simple graphs and $f: V(G) \to V(H)$ be a map. We say that f is isomorphism from G to H if it satisfies the following:

(a) f is one-to-one correspondence $i.e.$, f is bijective.

(b) $f: V(G) \to V(H)$ Preserves adjacency $i.e.$, for every $x, y \in V(G)$ then $\{x, y\} \in E(G)$ if and only if $\{f(x), f(y\} \in E(H)$ on the other words, x, y are adjacent in G if and only if $f(x), f(y)$ are adjacent in H . In this case we say that G and H are isomorphic and we write $G \cong H$.

Example.

Show that the following graphs $G = (V, E)$ and $H = (W, F)$ are isomorphic.



Solution.

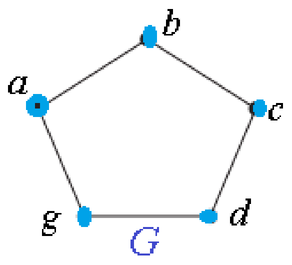
We define the mapping $f: V(G) \to V(H)$ as follows:

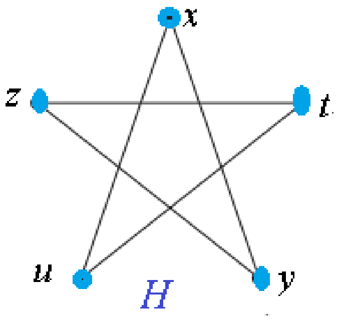
V	u_1	u_2	u_3	u_4
$f(V)$	v_1	v_2	v_4	v_3

It is obvious that f is one - to - one correspondence.To see that this correspondence preserves adjacency, note that adjacent vertices in G are $u_1$and u_2 , u_1 and u_3 , u_2 and u_4 , u_3 and u_4 , and each of the pair $f(u_1) = v_1$ and $f(u_2) = v_2$, $f(u_1) = v_1$ and $f(u_3) = v_4$, $f(u_2) = v_2$ and $f(u_4) = v_3$ and $f(u_3) = v_4$ and $f(u_4) = v_3$ are adjacent in H. Therefore the graphs $G = (V, E)$ and $H = (W, F)$ are isomorphic.■

Example.

Determine whether the following graphs are isomorphic or not? Explain your answer?





Solution.

We define the map $f: V(G) \to V(H)$ as follows:

v	a	b	c	d	g
$f(v)$	x	y	z	t	u

It easy to see that f is an isomorphism. Consequently $G \cong H$.■

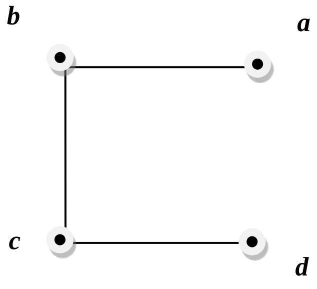
Note that in the above example H is the complement of G. So, we have the following definition:

Definition.

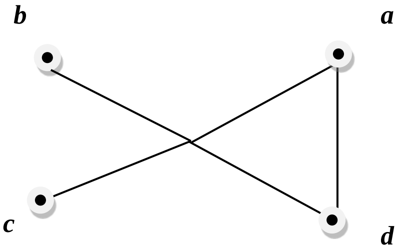
A simple graph G is said to be self-complementary if $G \cong \bar{G}$.

Example.

The following diagram is for a self-complementary Graph G (why:)



G



$\bar{G}$

It is often difficult to determine whether two simple graphs are isomorphic. However, we can use invariant with respect to isomorphism.

Definition.

We say that a property P is **invariant** with respect to isomorphism (isomorphism invariant) if the following condition is satisfied:

For every two simple graphs G and H, if $G \cong H$ and G has the property P, then H has the property P.

The following theorem gives us some isomorphism invariants. We can use them to discover non- isomorphic graphs.

Theorem 1. Let G and H be two simple graphs and $f: V(G) \to V(H)$ be an isomorphism**.** Then

(i) $|V(G)| = |V(H)|$ and $|E(G)| = |E(H)|$;

(ii) $d(x) = d(f(x))$ for every $x \in V(G)$;

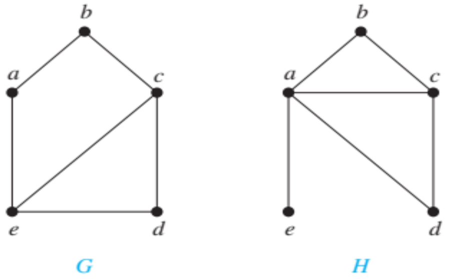
(iii) The number of vertices with degree m in G equals the number of vertices with degree m in H.

Proof : We accept (i) and (iii) , and prove only (ii) ,

(ii) Let $x \in V(G)$ with $d(x) = m$. Then there exist $x_1, \dots, x_m \in V(G)$ such that $x_i \neq x_j$ for every $i \neq j$ and x_i is adjacent to x for every i. Since f is one-to-one correspondence and preserves adjacency, then $f(x_1), \dots, f(x_m) \in V(H)$ are different vertices and each of them is adjacent to $f(x)$. Therefore $d\big(f(x)\big) \geq m$. Since f is subjective and preserves non- adjacency then the only vertices which are adjacency to the vertex $f(x)$ in H are $f(x_1), \dots, f(x_m)$. Therefore $d\big(f(x)\big) = m$. ■

Example.

Show that the graphs displayed in the figure are not isomorphic.

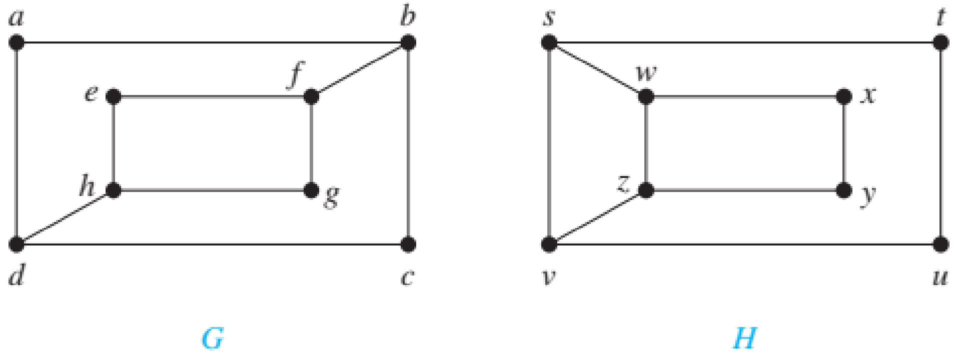


Solution.

Both G and H have five vertices and six edges. However, H has a vertex of degree one, namely, e, whereas G has no vertices of degree one. It follows that G and H are not isomorphic. ■

Example.

Determine whether the graphs shown in the following figure are isomorphic.



Solution.

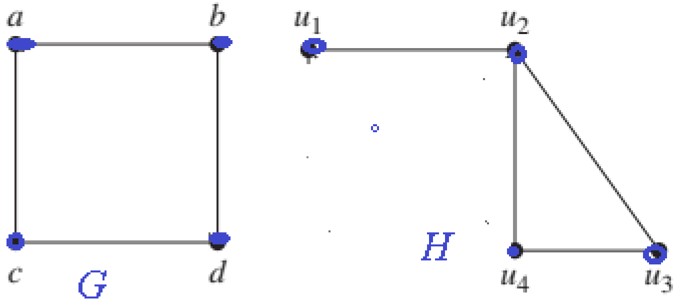
The graphs G and H both have 8 vertices and 10 edges. They also both have 4 vertices of degree 2 and 4 of degree 3. Because these invariants all agree, it is still conceivable that these graphs are isomorphic.

However, G and H are not isomorphic. To see this, note that because $d_G(a) = 2$, a must correspond to either $t, u, x,$ or y in H, because these are the vertices of degree two in H.

However, each of these four vertices in H is adjacent to another vertex of degree 2 in H, which is not true for a in G.

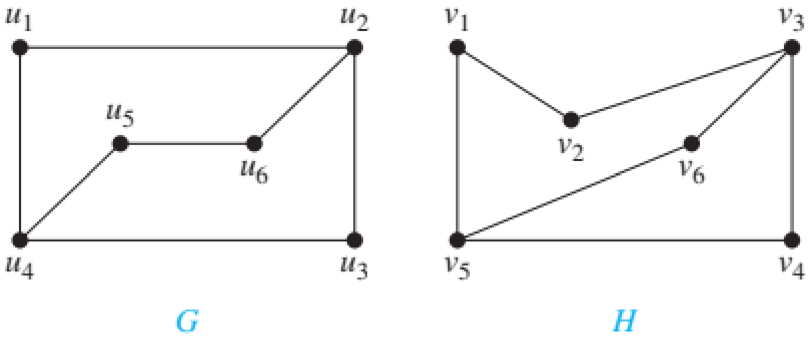
Example.

The following two graphs are not isomorphic because $d_H(u_2) = 3$ and there is no vertex in G with degree 3.



Example.

Determine whether the graphs G and H displayed in the following figure are isomorphic.



Solution.

Both G and H have six vertices and seven edges. Both have four vertices of degree two and two vertices of degree three. Because G and H agree with respect to these invariants, it is reasonable to find an isomorphism f.

We now will define a function f and then determine whether it is an isomorphism. Because $d_G(u_1) = 2$ and because u_1 is not adjacent to any other vertex of degree two, the image of u_1 must be either v_4 or v_6, the only vertices of degree two in H not adjacent to a vertex of degree two. We arbitrarily set $f(u_1) = v_6$. [If we found that this choice did not lead to isomorphism, we would then try $f(u_1) = v_4$.] Because u_2 is adjacent to u_1, the possible images of u_2 are v_3 and v_5. We arbitrarily set $f(u_2) = v_3$. Continuing in this way, using adjacency of vertices and degrees as a guide, we set $f(u_3) = v_4$, $f(u_4) = v_5$, $f(u_5) = v_1$, and $f(u_6) = v_2$. We now have a one-to-one correspondence between the vertex set of G and the vertex set of H. To see whether f preserves edges, we examine the adjacency matrix of G,

$$\mathbf{A}_G = \begin{array}{c} \\ u_1 \\ u_2 \\ u_3 \\ u_4 \\ u_5 \\ u_6 \end{array} \begin{array}{c} \begin{array}{cccccc} u_1 & u_2 & u_3 & u_4 & u_5 & u_6 \end{array} \\ \left[\begin{array}{cccccc} 0 & 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 0 & 1 & 0 \end{array}\right] \end{array},$$

and the adjacency matrix of H with the rows and columns labeled by the images of the corresponding vertices in G,

$$\mathbf{A}_H = \begin{array}{c} \\ v_6 \\ v_3 \\ v_4 \\ v_5 \\ v_1 \\ v_2 \end{array} \begin{array}{c} \begin{array}{cccccc} v_6 & v_3 & v_4 & v_5 & v_1 & v_2 \end{array} \\ \begin{bmatrix} 0 & 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 0 & 1 & 0 \end{bmatrix} \end{array}.$$

Because $AG = AH$, it follows that f preserves edges. We conclude that f is an isomorphism, so G and H are isomorphic. ■

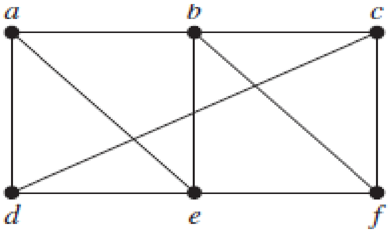
8.7 Connected Graphs

Definition.

Let $G = (V, E)$ be a graph, $a, b \in V$ and $n \geq 1$ be an integer. If $v_{1,}e_1, v_2, e_2, \dots.., e_{n-1}, v_n$ is a sequence of vertices and edges such that $v_1 = a, v_n = b, e_i = \{v_i, v_{i+1}\}$ for all i, then the sequence is called **a path** from a to b. **A path of length n** from a to b is a sequence of n edges. A path is **a circuit** (closed path) if $v_1 = v_n$. A path or circuit is **simple** : if $e_i \neq e_j$ for all $i \neq j, i.e.$, it does not contain the same edge more than one . When the graph is simple, we denote the path or the circuit by its vertices sequence $v_{1,}v_2, \dots.., v_n$. A simple circuit in which if $v_i \neq v_j$ for all $i \neq j$, except $v_1 = v_n$ is called **a cycle**.

Example 1.

In the simple graph.



1. $a\ , d, c, f$ is a simple path of length 3.

2. d, e, c, b is not a path since $\{e, c\}$ is not edge.

3. b, c, f, e, b is circuit of length 4. Also, it is a cycle.

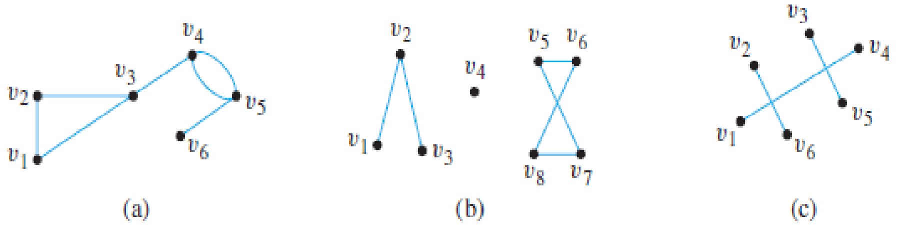
4.The path a, d, e, d, a, b which of length 5, is not simple since it contains the edge $\{a, d\}$ twice.

Definition.

An undirected graph is **connected** if there is a path between every pair of distinct vertices of the graph. We say it is **disconnected** if it is not connected.

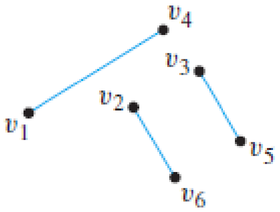
Example.

Which of the following graphs are connected?



Solution.

The graph represented in (a) is connected, whereas those of (b) and (c) are not. To understand why (c) is not connected, recall that in a drawing of a graph, two edges may cross at a point that is not a vertex. Thus, the graph in (c) can be redrawn as follows:

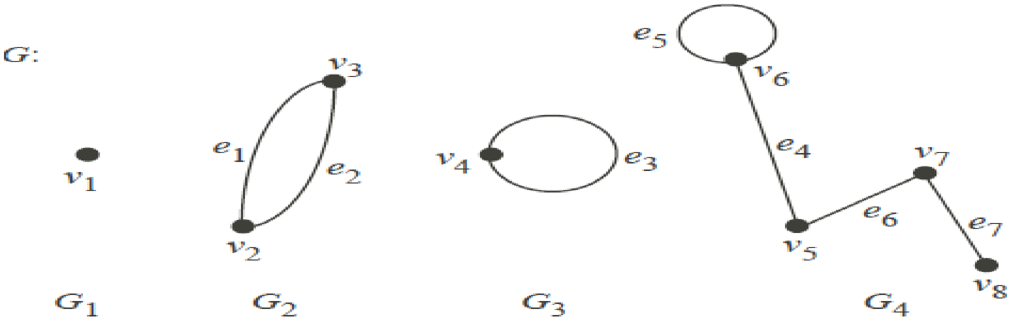


Theorem.

There is a simple path between every pair of distinct vertices of a connected undirected graph.

Definition.

A graph that is not connected is the union of two or more connected subgraphs, each pair of which has no vertex in common. These disjoint connected subgraphs are called the connected **components** of the graph.



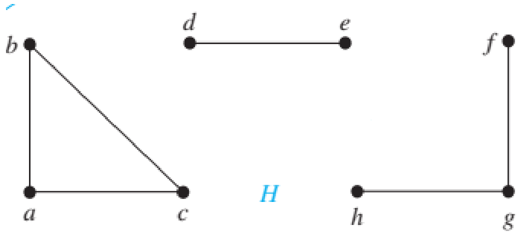
The components of G are G_1, G_2, G_3 and G_4.

Definition.

A **connected component** of a graph G is a connected subgraph of G that is not a proper subgraph of another connected subgraph of G. That is, a connected component of a graph G is a maximal connected subgraph of G. A graph G that is not connected has two or more connected components that are disjoint and have G as their union.

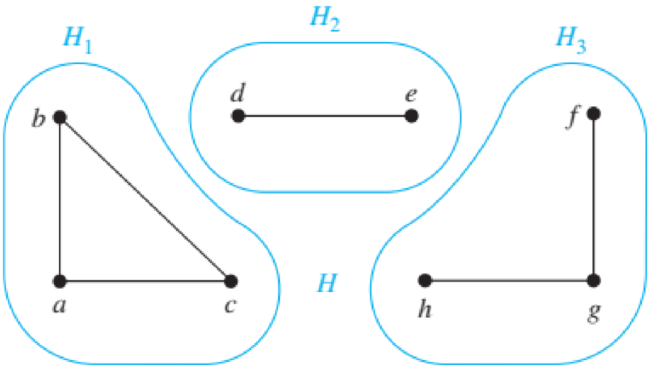
Example.

What are the connected components of given the graph H.



Solution.

The graph H is the union of three disjoint connected subgraphs H_1, H_2, and H_3, shown in the following figure. These three subgraphs are the connected components of H.



Theorem.

Each connected graph with n vertices has m edges where $m \geq n-1$.

Theorem.

In a simple graph G either G or $\bar{G}$ is connected.

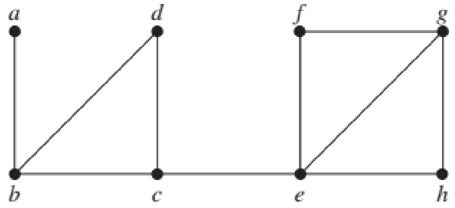
Definition.

Sometimes the removal of a vertex and all edges incident with it produces a subgraph with more connected components than in the original graph such vertices are called **cut vertices**. The removal of a cut vertex from a connected graph produces a subgraph that is not connected.

Analogously, an edge whose removal produces a graph with more connected components than the original graph is called a **cut edges** or **bridge**.

Example.

Find the cut vertices and cut edges in the following graph.



Solution.

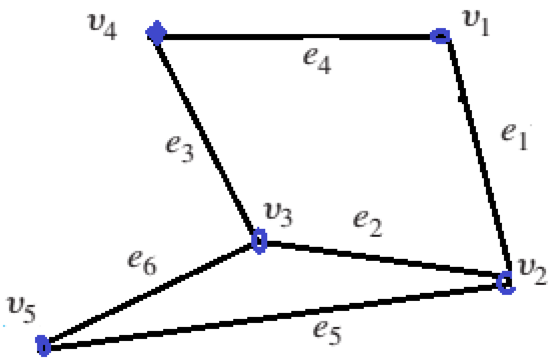
The cut vertices are $b, c,$ and e. The removal of one of these vertices (and its adjacent edges) disconnects the graph. The cut edges are $\{a, b\}$ and $\{c, e\}$. Removing either one of these edges disconnects. ■

Theorem.

Let $A = [a_{ij}]$ be the adjacency matrix for the graph $G = (V, E)$ such that $V = \{v_1, v_2, \dots, v_n\}$. Let $A^k = [b_{ij}]$ such that $k \geq 1$. Then the number of different paths from v_i to v_j with length k is equal b_{ij}.

Example.

Find the number of paths of length 4 from v_4 to v_5 for the following graph:



Solution.

The adjacency matrix is

$$A = \begin{bmatrix} 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 \end{bmatrix} \quad \text{and} \quad A^4 = \begin{bmatrix} 9 & 3 & 11 & 1 & 6 \\ 3 & 15 & 7 & 11 & 8 \\ 11 & 7 & 15 & 3 & 8 \\ 1 & 11 & 3 & 9 & 6 \\ 6 & 8 & 8 & 6 & 8 \end{bmatrix}$$

Hence $b_{45} = b_{54} = 6$.

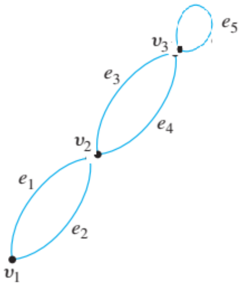
The number of paths from v_4 to v_5 is 6 and the paths are:

$v_4e_4e_4e_3e_6v_5, v_4e_3e_3e_3e_6v_5, v_4e_4e_1e_2e_6v_5, v_4e_3e_6e_6e_6v_5,$

$v_4e_3e_6e_5e_5v_5, v_4e_3e_2e_2e_6v_5.$■

Example.

Find the number of paths of length 3 from v_1 to v_3 and find the paths for the given graph.



Solution.

The adjacency matrix of the given graph is :

$$A = \begin{bmatrix} 0 & 2 & 0 \\ 2 & 0 & 2 \\ 0 & 2 & 1 \end{bmatrix}. \qquad \text{Then} \qquad A^3 = \begin{bmatrix} 0 & 16 & 4 \\ 16 & 4 & 18 \\ 4 & 18 & 9 \end{bmatrix}$$

The number of paths of length 3 between v_1 and v_3 is 4:

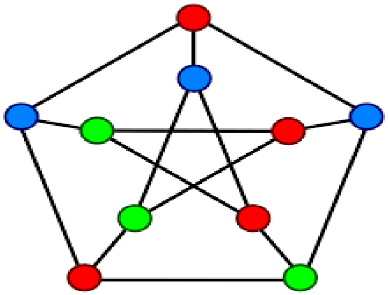
The paths are

$x_1e_1e_3e_5x_3, x_1e_1e_4e_5x_3, x_1e_2e_3e_5x_3, x_1e_2e_4e_5x_3.$■

Theorem.

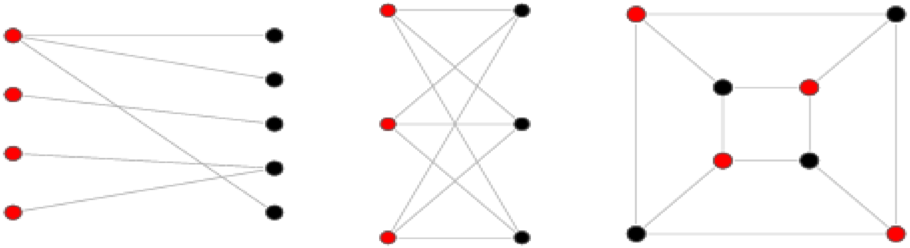
The simple graph G is bipartite graph if G has no odd cycle.

Example.



Petersen graph above conations cycle with length 5. So, it is not bipartite graph. ■

Example.



The above graphs contain no odd cycles. Therefore, they are bipartite graphs. ■

8.8 Planar graph

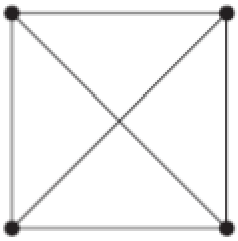
Definition.

A graph is called **planar** if it can be drawn in the plane without any edges crossing.

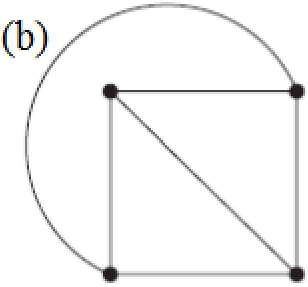
Example.

Although the complete graph K_4 is usually pictured with crossing edges as in figure (a), it can also be drawn with no crossing edges as in figure (b). Thus k_4 is a planar graph. Also, Q_3 in (c) and (d). Therefore, Q_3 is a planar graph. Such a drawing is called a *planar representation* of the graph.

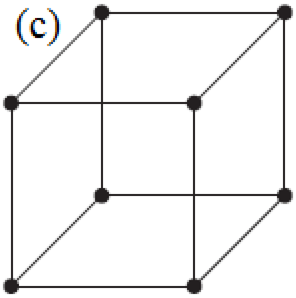
(a)



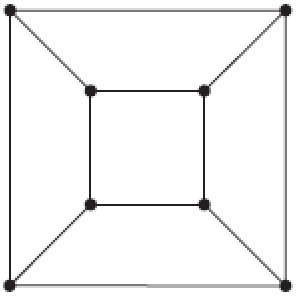
(b)



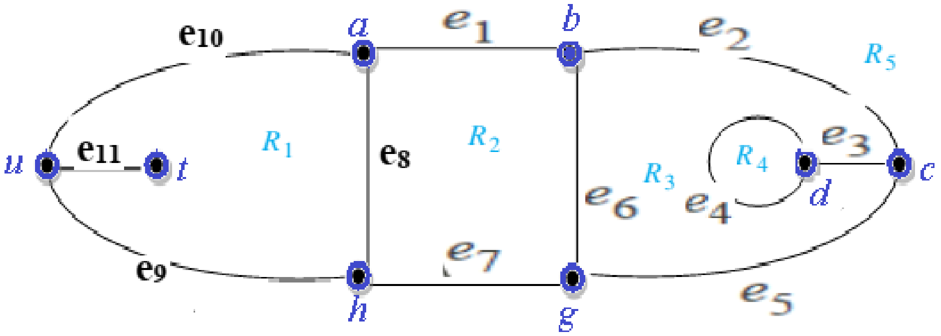
(c)



(d)



A planar representation of a graph divides the plane. For instance, consider the connected planar graph given in the following Figure



It is clear that the graph divides the plane into five regions. All regions are bordered except R_5.

The border of each one is as follows

R_1 bordered by the closed path $ue_{10}ae_8he_9ue_{11}te_{11}u$.

R_2 is bordered by the cycle $ae_1be_6ge_7he_8a$.

R_3 is bordered by the closed path $be_2ce_3de_4de_3ce_5ge_6b$.

R_4 is bordered by de_4d.

R_5 is bordered inside by the cycle $ae_1be_2ce_5ge_7he_9ue_{10}a$.

By the degree of a region R, written $d(R)$, we mean the length of the cycle or closed simple path border R. ■

Theorem.

The sum of degrees of the regions of a planar representation of a graph is twice the number of edges. We note that each edge either borders two regions or is contained in a region and will occur twice in any simple path along the border of the region.

Example.

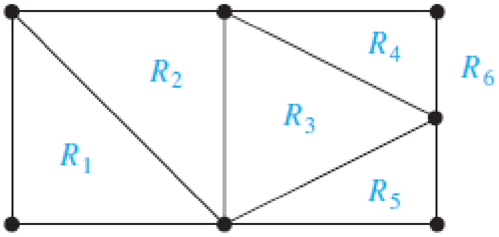
The degrees of the regions of the above figure are $d(R_1) = 5, d(R_2) = 4, d(R_3) = 6, d(R_4) = 1, d(R_5) = 6$.

The sum of the degrees is 22 which is twice the number of edges, as expected.■

The Regions of the Planar Representation of a Graph.

♣ Euler's formula.

A planar representation of a graph splits the plane into regions, including an unbounded region.
For instance, the planar representation of the graph shown in the following figure splits the plane into six regions.



Euler's formula connects the number of vertices v , the number of edges e and the number of regions r of any connected simple planar graph G. Euler's formula is:

$$v - e + r = 2$$

Euler's formula is special for **connected simple planar graph.** If G is a planar graph with K components, then one can deduce that: $v - e + r = K + 1$.

Example.

Suppose that a connected planar simple graph has 20 vertices, each of degree 3. How many regions does a representation of this planar graph split the plane?

Solution.

This graph has 20 vertices, each of degree 3, so $v = 20$. Because the sum of the degrees of the vertices, $3v = 3 \cdot 20 = 60$, is equal to twice the number of edges, $2e$, we have $2e = 60$, or $e = 30$. Consequently, from Euler’s formula, the number of regions is

$$r = e - v + 2 = 30 - 20 + 2 = 12. ■$$

Theorem.

(a) Let G be a connected planar simple graph such that $v \geq 3$, then $e \leq 3v - 6$;

(b) K_5 is not a planar graph.

Proof.

Since G is connected and $v \geq 3$, we have $e \geq 2$. Hence $2 \leq 3(3) - 6 = 3$ and the inequality is true. Now, suppose $e \geq 3$. Since the Sam of the degrees of the regions is $2e$. But each region has degree three or more. Because at least 3 edges border one region. Therefore, $3r \leq 2e$. But from Euler's formula $v - e + r = 2$. Then $3[2 - v + e] = 3r \leq 2e$. Therefore $e \leq 3v - 6$.

(b) Suppose that K_5 is a planar graph. We know that $v = 5$, $e = 10$. Since K_5 is a simple connected graph, then from (a) above we have $10 \leq 3(5) - 6 = 9$, which is a contradiction. Therefore K_5 is not a planar graph.■

Theorem.

(a) Let G be a connected planar simple graph with $v \geq 3$ and G has no cycle of length 3. Then $e \leq 2v - 4$.

(b) $K_{3,3}$ is not a planar graph.

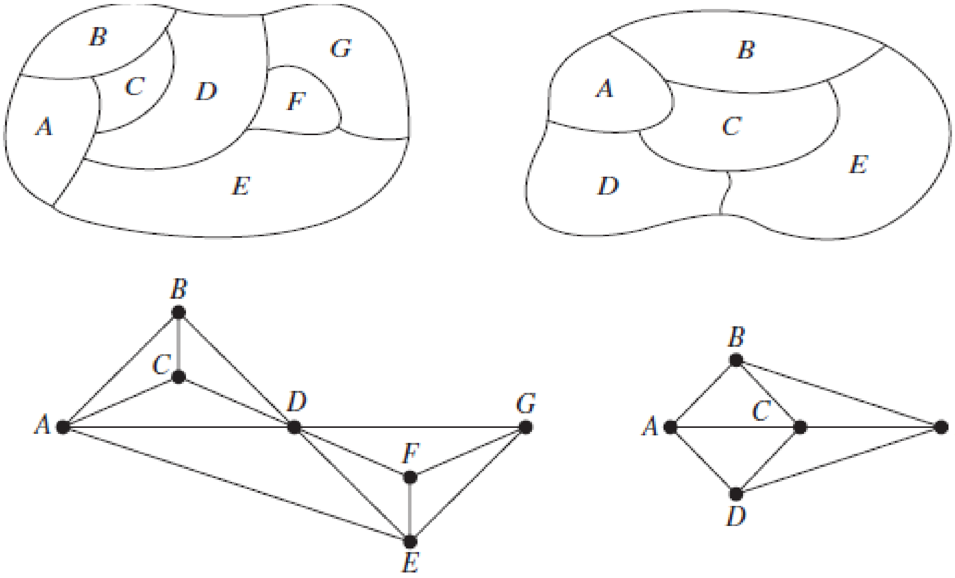
Proof.

(a) Since G is connected and $v \geq 3$, then $e \geq 2$. Since the sum of degrees of the regions is $2e$. But each region has degree 4 or more because G has no cycle of length 3 , *i,e*, at least 4 edges border one region. Hence $4r \leq 2e$. By Euler's formula we have $r = 2 - v + e$. Hence $4[2 - v + e] \leq 2e$. Therefore $e \leq 2v - 4$.

(b) Suppose $K_{3,3}$ is a planar graph. We know that $v = 6$, $e = 9$. Since $K_{3,3}$ is a connected planar simple graph and has no cycle of length 3, then $9 \leq 2 \times 6 - 4 = 8$, a contradiction. Therefore $K_{3,3}$ is not a planar graph.■

♣ Graph coloring

Each map in the plane can be represented by a planar a graph, where each region of the map is represented by a vertex. Edges connect two vertices if the region represented by these vertices has a common border. Two regions that touch at only one point are not considered adjacent. The resulting graph is called the **dual** graph of the map. Let M be a map and $G = (V, E)$ be a planar graph represents the map M (the dual graph). Where V is the regions in the map and $\{x, y\} \in E$ if and only if the two regions x and y are adjacent:



The dual graphs of the given maps.

Definition.

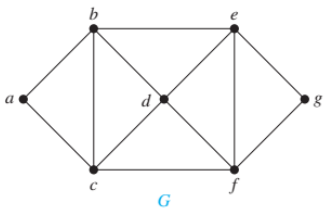
A **coloring** of a simple graph is the assignment of a color to each vertex of the graph so that no two adjacent vertices are assigned the same color. The least number of colors need to color the graph G is called chromatic number $\chi(G)$.

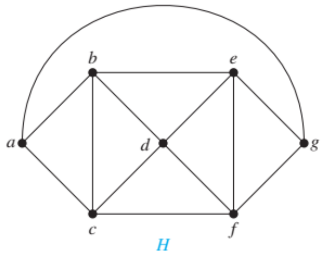
Four colors problem:

" Is it possible to color a map with at most 4 colors so that no adjacent regions have the same color" **or equivalently** "If G is a simple planar graph, then $\chi(G) \leq 4$ "?

Example.

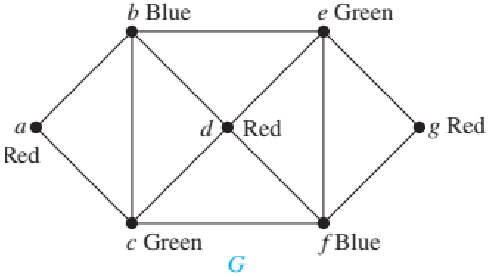
What are the chromatic numbers of the graphs G and H shown in the following figure?



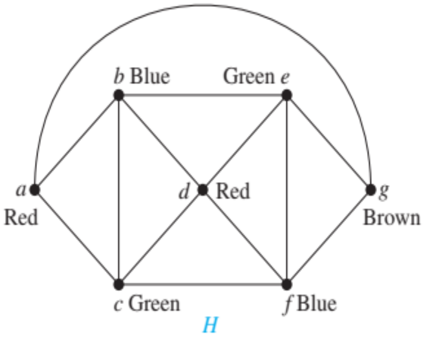


Solution.

The chromatic number of G is at least three, because the vertices a, b, and c must be assigned different colors. To see if G can be colored with three colors, assign red to a, blue to b, and green to c. Then, d can (and must) be colored red because it is adjacent to b and c. Furthermore, e can (and must) be colored green because it is adjacent only to vertices colored red and blue, and f can (and must) be colored blue because it is adjacent only to vertices colored red and green. Finally, g can (and must) be colored red because it is adjacent only to vertices colored blue and green. This produces a coloring of G using exactly three colors. The given figure displays such a coloring.



The graph H is made up of the graph G with an edge connecting a and g. Any attempt to color H using three colors must follow the same reasoning as that used to color G, except at the last stage, when all vertices other than g have been colored. Then, because g is adjacent (in H) to vertices colored red, blue, and green, a fourth color, say brown, needs to be

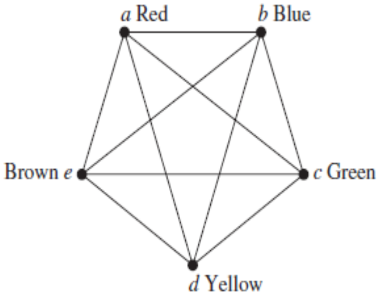


used. Hence, H has a chromatic number equal to 4. A coloring of H is shown in the given figure. ■

Example.

It is clear that if $G = (V, E)$ is a simple graph with $|V| = n$, then $\chi(G) \leq n$. If H is a subgraph of G. then $\chi(H) \leq \chi(G)$.

In K_n, $d(x) = n - 1$ for every $x \in V$. Therefore $\chi(K_n) = n$ and $\chi(\overline{K_n}) = 1$. For example, K_5. ■

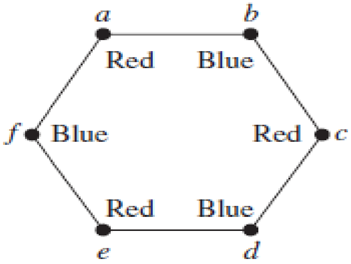


Example.

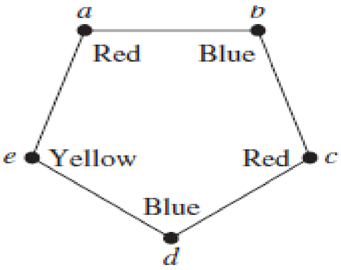
If C_n is a cycle with length n, then $\chi(C_n) = 2$ if n is even and $\chi(C_n) = 3$ if n is odd.

Solution.

let $x_1, \dots, x_n$ be the vertices of C_n. If we colored x_1 by color 1, then x_2 should take different color (say color 2). Thus x_3 can take color 1. So, if n is odd then we need a third color 3 to color x_n , but, if n is even, then x_n colored by color 2. ■



C_6



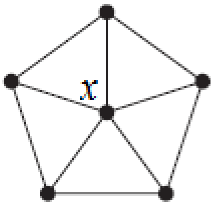
C_5

Example.

Calculate $\chi(W_5)$, when W_5 is the wheel graph shown the following diagram.

Solution.

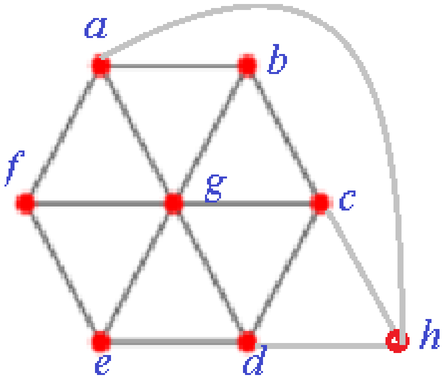
Since $W_5 - x \cong C_5$, then we need 3 colors to color C_5. Since x is adjacent to all vertices in W_5 , then we need another color to color it. Consequently, $\chi(W_5) = 4$. ■



W_5

Example.

Calculate $\chi(G)$, for the given graph.



G

Solution.

Note that $G-\{g,h\}\cong C_6$ and $\chi(C_6)=2$. But g is adjacent to all vertices of C_6, and then we need a third color (3) for g. Hence, we can choose color (3) to h.

Therefore $\chi(G)=3$.■

Theorem.

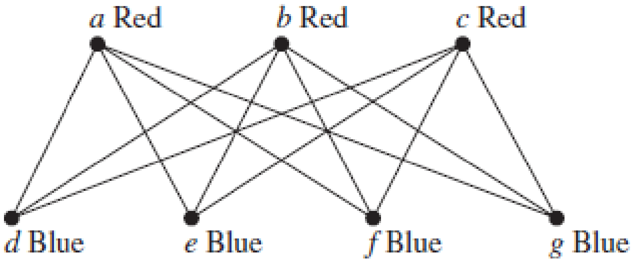
$\chi(G)=2$ if and only if G is a bipartite graph.

Proof.

Let G be a bipartite graph and $V=V_1\cup V_2$. It is enough to color V_1 by only one color and V_2 by another one. So $\chi(G)=2$.

Conversely, suppose $\chi(G)=2$ and V_1 be the set colored by first color and V_2 be the set colored by the second one. There is no adjacent two vertices of $V_1(or\ V_2)$. So any edge in G should connect a vertex in V_1 and a vertex in V_2. Hence G is bipartite graph.■

K$_{3,4}$:



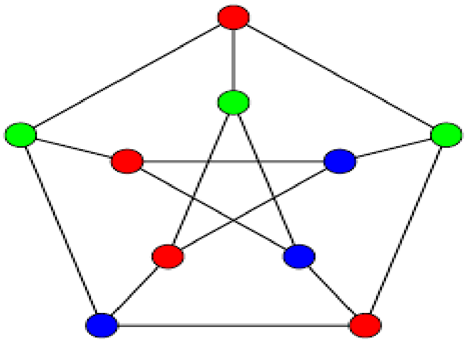
Corollary.

$\chi(G) \geq 3$ if and only if G contains an odd cycle.

Example.

Show that $\chi(G) = 3$, where G is Petersen graph.

Solution.



Since G contains a cycle of length 5, then $\chi(G) \geq 3$. Only 3 colors are enough as shown in the Figure. ■

Graph coloring has a variety of applications to problems involving scheduling and assignments. The following example is one of these applications.

Example.

How can the final exams at a university be scheduled so that no student has two exams at the same time?

Solution.

This scheduling problem can be solved using a graph model, with vertices representing courses and with an edge between two vertices if there is a common student in the courses they represent. Each time slot for a final exam is represented by a different color. A scheduling of the exams corresponds to a coloring of the associated graph. ■

See the following example

Example.

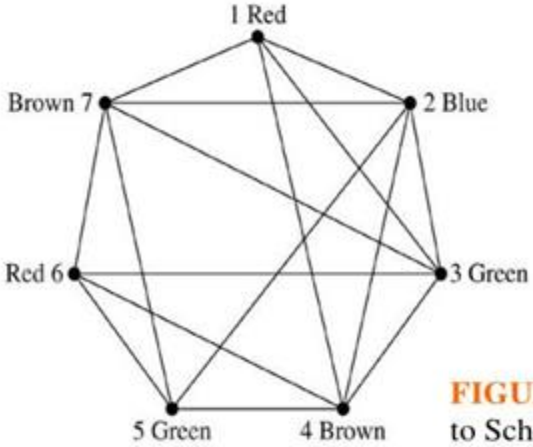
Suppose there are seven finals to be scheduled. Suppose the courses are numbered 1 through 7. Suppose that there is one student or more scheduled in each of: $(1,2)$, $(1,3)$, $(1,4)$, $(1,7)$, $(2,3)$, $(2,4)$, $(2,5)$, $(2,7)$, $(3,4)$, $(3,6)$, $(3,7)$, $(4,5)$, $(4,6)$, $(5,6)$, $(5,7)$, $(6,7)$.

Use graph coloring to schedule the final exams so that no student has two exams at the same time and we have the least time slots.

Solution.

Let G be the graph with vertices representing courses and $\{x,y\} \in E$ if and only if a student or more are scheduled in the courses x and y. The following diagrams shows the coloring graph.

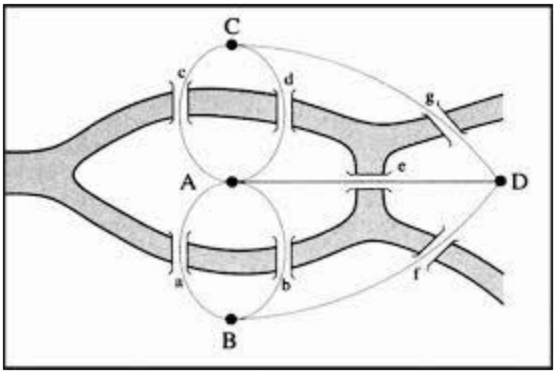
A scheduling consists of a coloring of this graph. We need four colors to color this graph $(\chi(G) = 4)$. So we need 4 time slots as shown in the table:



Time Period	Courses
I	1, 6
II	2
III	3, 5
IV	4, 7

FIGURE Using a Coloring to Schedule Final Exams.

♣ Euler paths

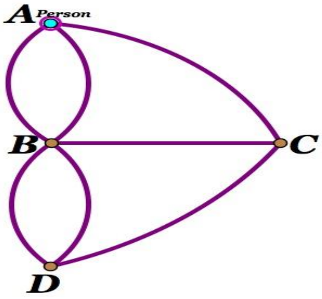


Pregl River divide Konigsberg city in Germany into 4 parts included two islands A and D and two regions B and C. There are seven bridges connected these sections.

The **seven bridges problem** say that:

" Is it possible to start at some location in the town travel across all the bridges without crossing any bridge twice, and return to the starting point?"

The Swiss mathematician Euler studied this problem using the multi-graph obtained when the four regions are represented by vertices and bridges by edges.



The question become : Is there a simple circuit in this multi-graph that contains every edge? The answer was no.

Definition.

1. An **Euler circuit** in a graph G is a simple circuit containing every edge in G. In this case G is called **Euler graph**.

2. An Euler path in G is a simple path containing every edge in G. In this case G is called **half Euler graph**.

Theorem.

A connected multi-graph has an Euler circuit if and only if each of its vertices has an even degree.

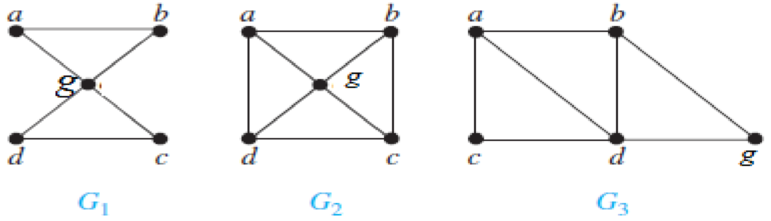
Theorem.

A connected multi-graph has an Euler path if and only if it has only two odd vertices.

We can now solve the seven bridges problem. Since the multi-graph representing these bridges has four vertices of odd degree, it does not have an Euler circuit.

Example.

Which of the undirected graphs in the figure have an Euler circuit? Which have an Euler path?



Solution.

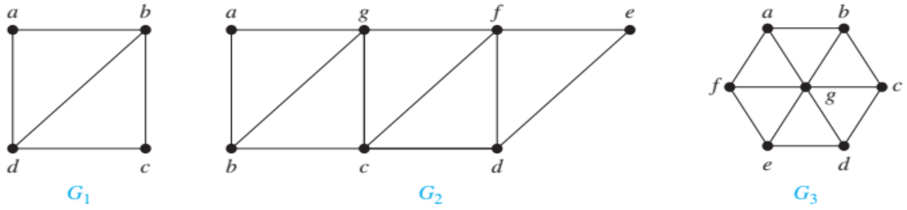
The graph G_1 has an Euler circuit. (Connected graph with even vertices), for example $a\ , g\ , c\ , d\ , g\ , b\ , a$.

Neither G_2 nor G_3 has an Euler circuit. However, G_3 has an Euler path (connected graph with two odd vertices), namely $a\ , c\ , d, g\ , b, d\ , a\ , b$.

G_2 does not have an Euler path. ■

Example.

Which graphs shown in the following figure have an Euler path?



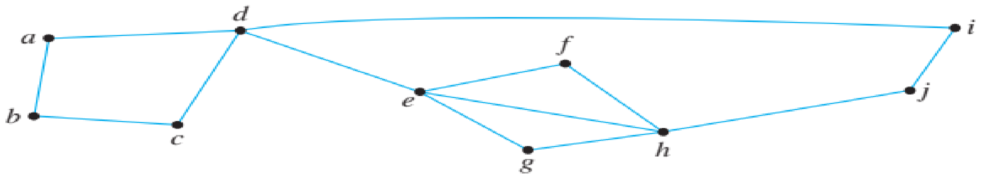
Solution.

G_1 contains exactly two vertices of odd degree, namely, b and d. Hence, it has an Euler path that must have b and d as its endpoints. One such Euler path is d, a, b, c, d, b.

Similarly, G_2 has exactly two vertices of odd degree, namely, b and d. So it has an Euler path that must have b and d as endpoints. One such Euler path is $b, a, g, f, e, d, c, g, b, c, f, d$. G_3 has no Euler path because it has six vertices of odd degree. ■

Example.

Is the following graph shown in the following figure have an Euler circuit? If yes find it.

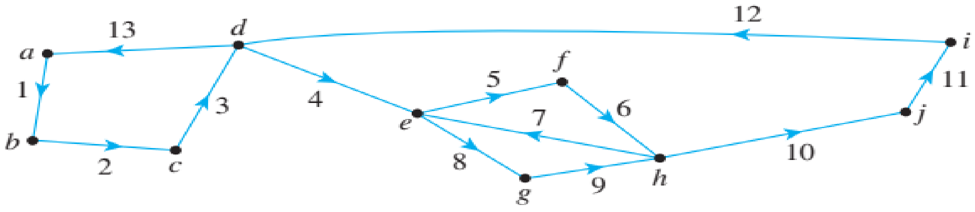


Solution.

Hence all vertices have even degree. Also, the graph is connected. Thus, the graph has an Euler circuit. It is:

$$a, b, c, d, e, f, h, e, g, h, j, i, d, a.$$

Then the Euler circuit is represented by the labeled edges shown below as it includes every edge of the graph exact once.■



Exercises Set (8)

1. Let $G = (V, E, f)$ be a graph, where $V = \{a, b, c\}, E = \{e_1, e_2, e_3, e_4\}$ and f represented by the following table:

e	e_1	e_2	e_3	e_4
f(e)	$\{a , b\}$	$\{a , b\}$	$\{a , b\}$	$\{b, c\}$

a. Represent the graph G ;

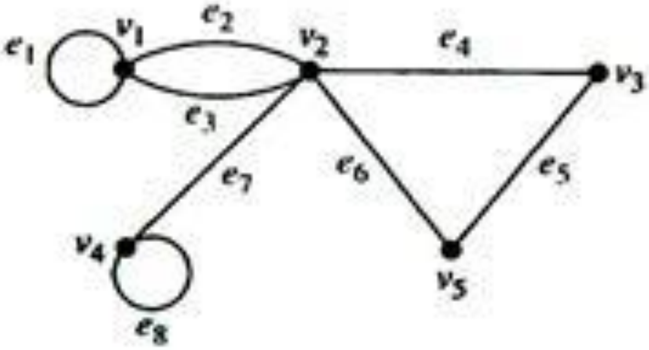
b. Find the degrees of vertices ;

c. Find loops and multi-edges ;

d. Is G simple graph ?Why ;

e. Find the adjacency matrix and the incidence matrix of G.

2. Find f , E , V, where $G = (V, E, f)$ is :



3. Give an example of a simple graph with odd vertices and other with even vertices.

4. Find the number of edges of a graph $G = (V, E)$, where the sum of its vertices degrees is 48.

5. Is there a graph the sequence of vertices degrees is :

a. $5, 5, 5, 3, 2, 2, 1$; (b) $3, 3, 3, 3, 3$.

6. Let $A_1, A_2, \dots., A_n$ be sets. The graph of their intersections is the simple graph whose vertices is $A_1, A_2, \dots., A_n$ and A_i is adjacent to A_j if $A_i \cap A_j \neq \emptyset$.

Find the graph if :

a. $A_1 = \{0,2,4,6,8\}, A_2 = \{0,1,2,3,4\} \quad A_3 = \{1,3,5,7,9\}$, $A_4, \{5,6,7,8,9\}$, $A_5 = \{0,1,8,9\}$.

b. $A_1 = (-\infty, 0), A_2 = (-1,0), A_3 = (0,1)$,

$A_4 = (-1,1), A_5 = (-1, \infty), A_6 = R$

7. Represent $K_6, K_7, K_{1,8}, K_{4,4}$.

8. Find the number of edges of a graph, the sequence of its vertices is 2, 2, 3, 3, 4 and represent it.

9. Is there a simple graph the sequence of its vertices $6, 4, 3, 2, 2, 1$.

10. Is there a graph the number of its vertices is 10 and the number of its edges is 50?

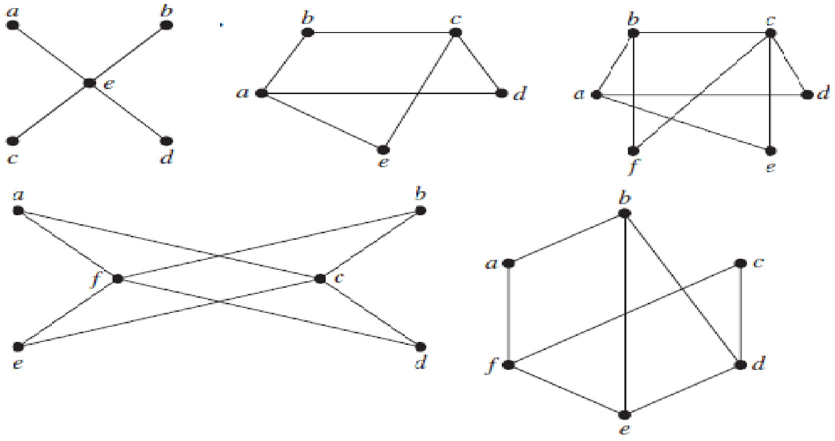
11. Give an example of 2-regular bipartite graph with 6 vertices

12. Give an example of 3-regular bipartite graph with 8 vertices

13. Suppose that $G = (V, E)$ is a simple graph with n vertices and $|E| > \frac{n^2}{4}$. Prove that G is not complete bipartite graph.

14. Let $G = (V, E)$ be a simple 4-regular graph with 10 edges. Compute its vertices.

15. Determine whether the following graph is bipartite graph or not? Give a suitable partition for each bipartite one.



16. Schedule the final exams for Math115, Math116, Math185, Math195, CS101, CS102, CS273 and CS473 using the fewest number of different time slots, if there are no students taking both Math115 and CS473, both

Math116 and CS473, both Math195 and CS101, both Math195 and CS102, both Math115 and Math116, both Math115 and Math185 and both Math185 and Math195 but there are student in every other combination of courses.

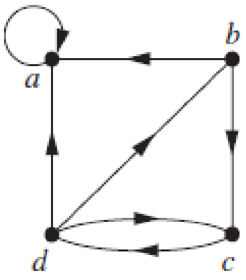
17. Show that a simple graph with a chromatic number of 2 is bipartite.

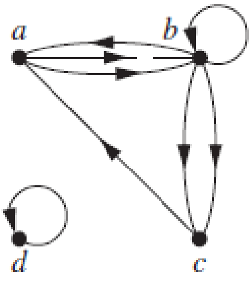
18. Show that a connected bipartite graph has a chromatic number of 2.

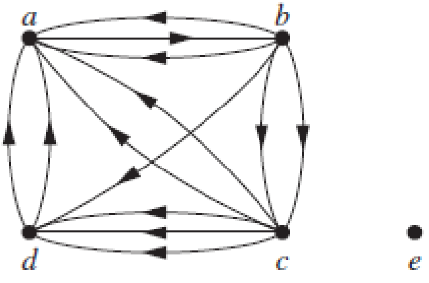
19. Show that $m \leq 2n - 4$ for a planar bipartite graph of n vertices and m.

20. Show that every planar graph contains a vertex of degree at most five.

21. Determine the number of vertices and edges and find the in-degree and out-degree of each vertex and find the adjacency matrix for the given directed multigraph.







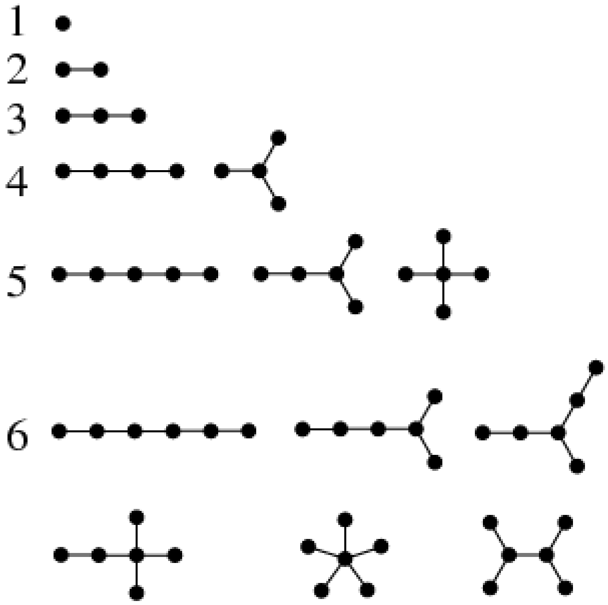
CHAPTER (IX)

TGEES

CHAPTER (IX)

Trees

9.1 Tree



In mathematics, and more specifically in graph theory, a **tree** is an undirected graph in which any two vertices are connected by *exactly one* simple path. In other words, any connected graph without simple cycles is a tree. The various kinds of data structures referred to as trees in computer science are equivalent as undirected graphs to trees in graph theory, although such data structures are generally **rooted trees**, thus in fact being directed graphs, and may also have additional ordering of branches.

Definition.

● A **tree** is an undirected simple graph G that satisfies any of the following equivalent conditions:

1. G is connected and has no cycles.

2. G has no cycles, and a simple cycle is formed if any edge is added to G.

3. G is connected, but is not connected if any single edge is removed from G.

4. G is connected and the 3-vertex complete graph K_3 is not a minor of G.

5. Any two vertices in G can be connected by a unique simple path.

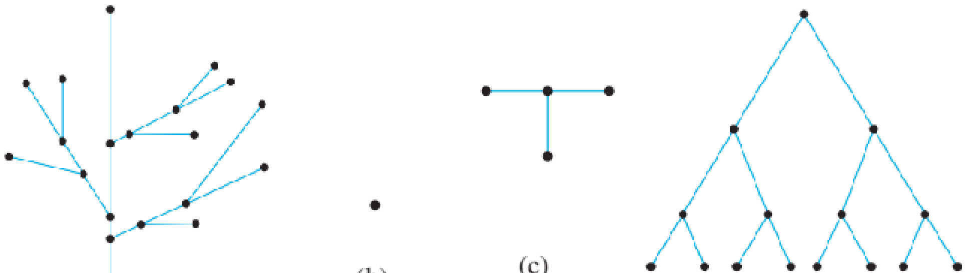
If G has finitely many vertices, say n of them, then the above statements are also equivalent to any of the following conditions:

6. G is connected and has $n - 1$ edges.

7. G has no simple cycles and has $n - 1$ edges.

Example.

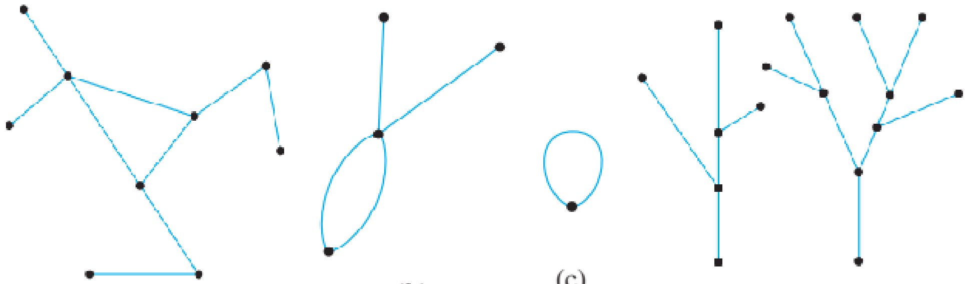
All the graphs shown in the following figure are trees.



(a) (b) (c) (d)

Example.

All the graphs shown in the following figure are not trees.

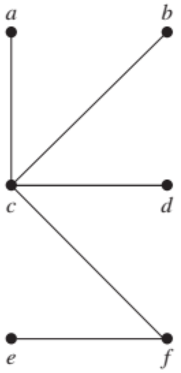


(a) (b) (c) (d)

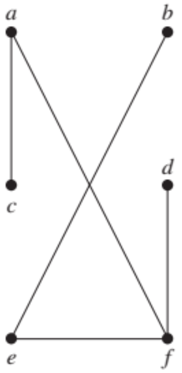
The graphs in (a), (b), and (c) all have circuits, and the graph in (d) is not connected. ■

Example.

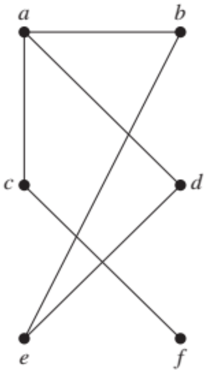
Which of the graphs shown in the following figure are trees?



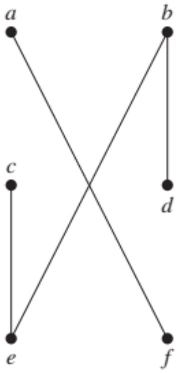
G_1



G_2



G_3



G_4

Solution.

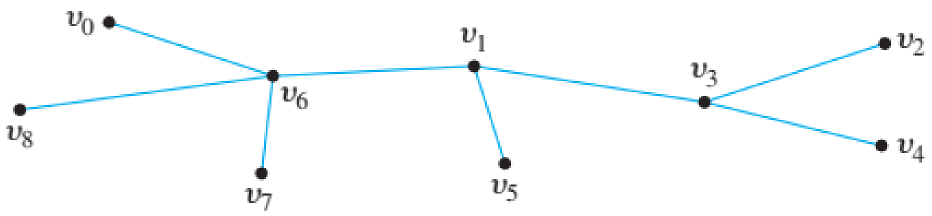
G_1 and G_2 are trees, because both are connected graphs with no simple circuits. G_3 is not a tree as e, b, a, d, e is a simple circuit in this graph. Finally, G_4 is not a tree because it is not connected. ■

Definition.

Let T be a tree. If T has at least two vertices, then a vertex of degree 1 in T is called a **leaf** (or a **terminal vertex**), and a vertex of degree greater than 1 in T is called an **internal vertex** (or **a branch vertex**). The unique vertex in a trivial tree is also called a leaf or terminal vertex.

Example.

Find all leaves (or terminal vertices) and all internal (or branch) vertices in the following tree:



Solution.

The leaves (or terminal vertices) are v_0, v_2, v_4, v_5, v_7, and v_8. The internal (or branch) vertices are v_1, v_3, and v_6. ■

Example.

A graph G has ten vertices and twelve edges. Is it a tree?

Solution.

No. Since any tree with n vertices has $n - 1$ edges, then any tree with ten vertices has nine edges, not twelve. ■

Theorem.

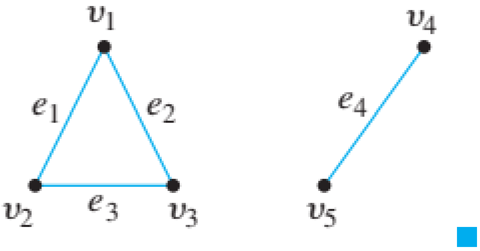
For any positive integer n, if G is a connected graph with n vertices and $n - 1$ edges, then G is a tree.

Example.

Give an example of a graph with five vertices and four edges that is not a tree.

Solution.

By the above theorem, such a graph cannot be connected. One example of such an unconnected graph is shown below.



9.2 Examples of Trees

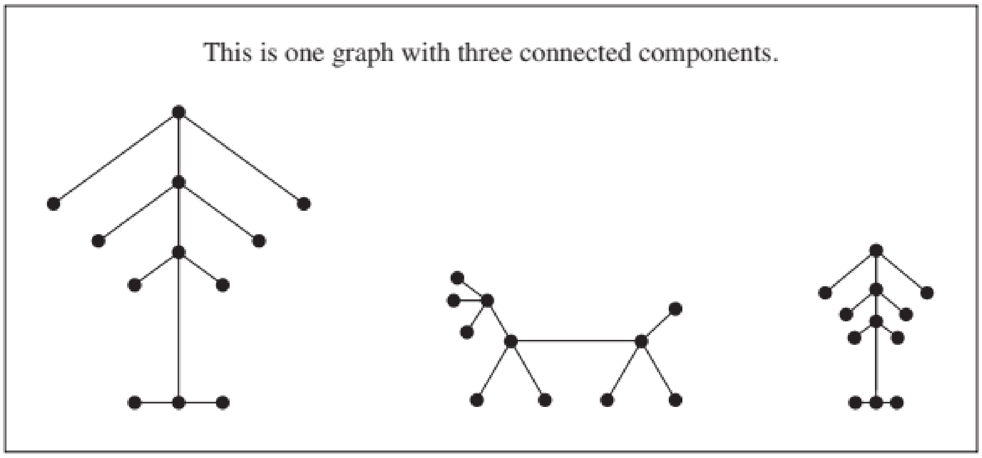
●Forest

Definition.

A **forest** is an undirected graph, all of whose connected components are trees; in other words, the graph consists of a disjoint union of trees. Equivalently, a forest is an undirected cycle-free graph. As special cases, an empty graph, a single tree, and the discrete graph on a set of vertices (that is, the graph with these vertices that has no edges), all are examples of forests.

Example.

The following figure displays a forest.

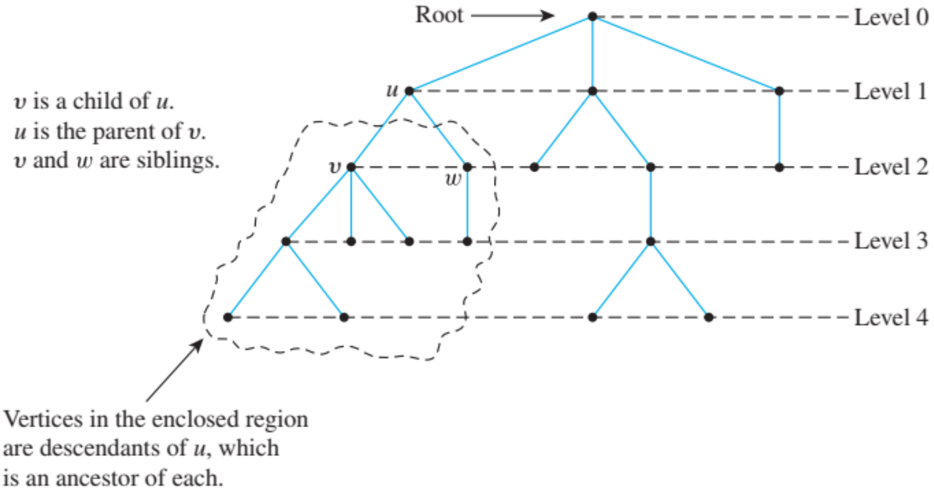


●A rooted tree

Definition.

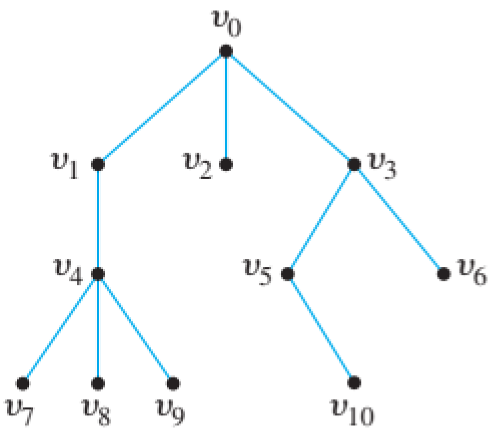
A **rooted tree** is a tree in which there is one vertex that is distinguished from the others and is called the **root**. The **level** of a vertex is the number of edges along the unique path between it and the root. The **height** of a rooted tree is the maximum level of any vertex of the tree. Given the root or any internal vertex v of a rooted tree, the **children** of v are all those vertices that are adjacent to v and are one level farther away from the root than v. If w is a child of v, then v is called the **parent** of w, and two distinct vertices that are both children of the same parent are called **siblings**. Given two distinct vertices v and w, if v lies on the unique path between w and the root, then v is an **ancestor** of w and w is a **descendant** of v.

These terms are illustrated in the following figure.



Example.

Consider the tree with root v_0 shown below.



a. What is the level of v_5?

b. What is the level of v_0?

c. What is the height of this rooted tree?

d. What are the children of v_3?

e. What is the parent of v_2?

f. What are the siblings of v_8?

g. What are the descendants of v_3?

h. How many leaves (terminal vertices) are on the tree?

Solution.

a. 2

b. 0

c. 3

d. v_5 and v_6

e. v_0

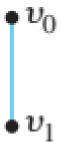
f. v_7 and v_9

g. v_5, v_6 and v_{10}

h. 6.■

Example.

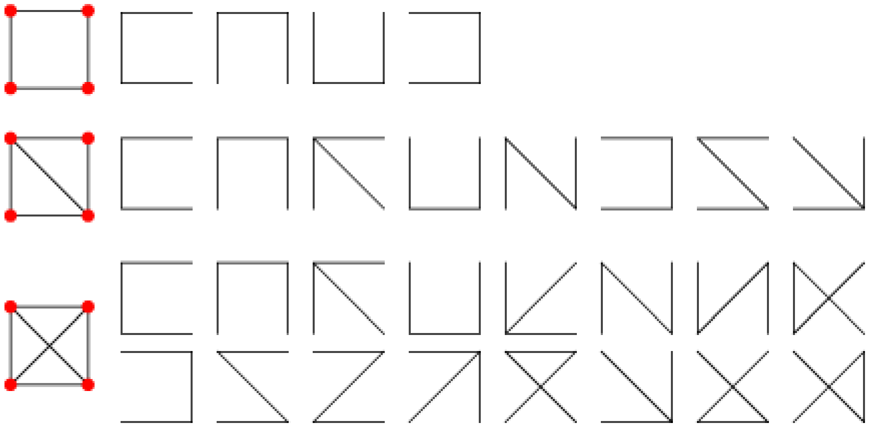
In the given tree, the root is v_0, v_1 has level 1, v_1 is the child of v_0, and both v_0 and v_1 are leaves (terminal vertices).



● Spanning Tree

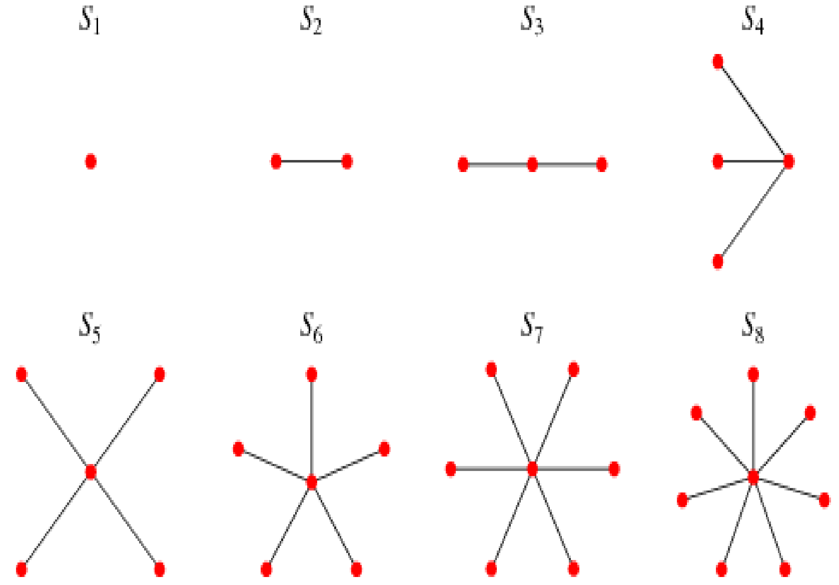
Definition.

A spanning tree of a graph on n vertices is a subset of $n-1$ edges that form a tree. For example, the spanning trees of the cycle graph C_4, diamond graph, and complete graph K_4 are illustrated above.



The number of non-identical spanning trees of a graph G is equal to any cofactor of the degree matrix of G minus the adjacency matrix of G. This result is known as the matrix tree theorem. A tree contains a unique spanning tree, a cycle graph C_n contains n spanning trees, and a complete graph K_n contains n^{n-2} spanning trees.

● Star Graph

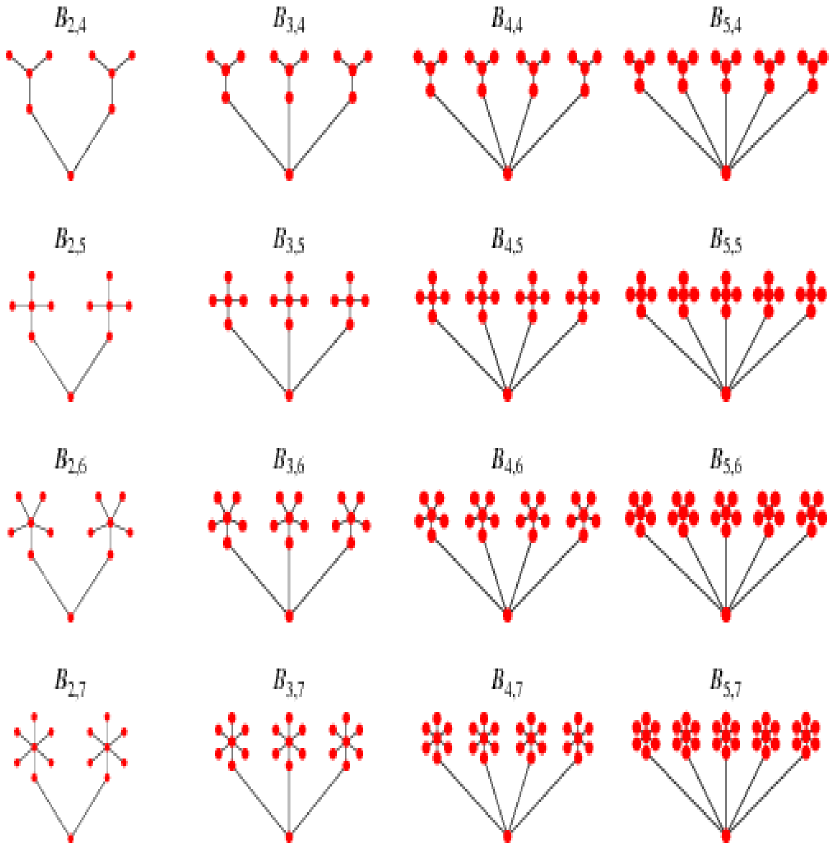


The star graph S_n of order n, sometimes simply known as an "n-star" is a tree on n nodes with one node having vertex degree $n-1$ and the other $n-1$ having vertex degree 1. The star graph S_n is therefore isomorphic to the complete bipartite graph $K_{1,n-1}$. The chromatic number is 1 for $n=1$, and $\chi(S_n)=2$ otherwise.

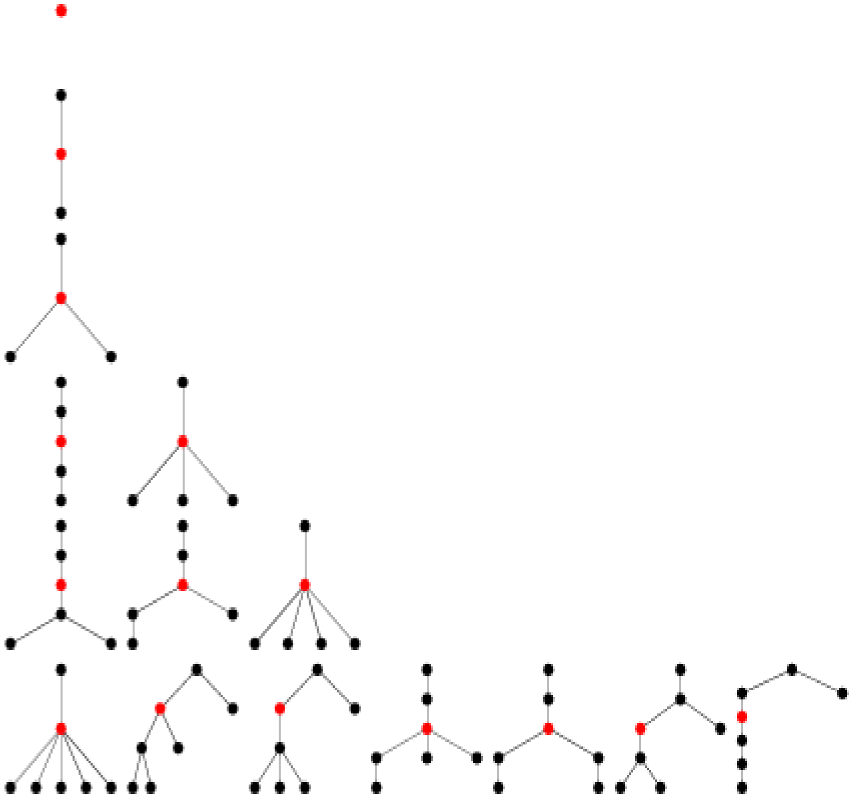
● Banana Tree

Definition.

An (n, k) -banana tree is a graph obtained by connecting one leaf of each of n copies of an k-star graph with a single root vertex that is distinct from all the stars.

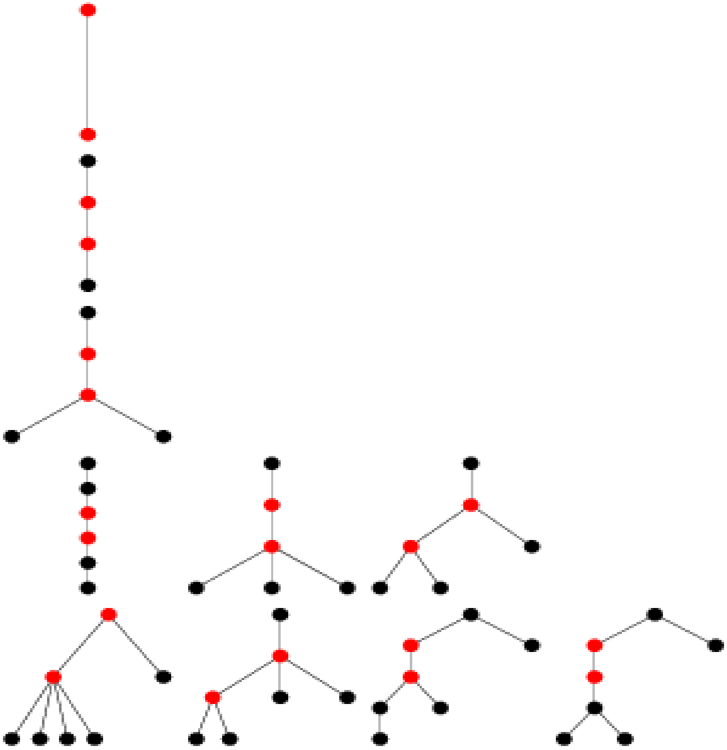


● Centered Tree



A tree (also called a central tree) having a single node that is a graph center. The numbers of centered trees on $n = 1,2,\ldots$ nodes are 1, 0, 1, 1, 2, 3, 7, 12, 27, 55, 127, 284, 682, 1618, ...

● Bi-centered Tree

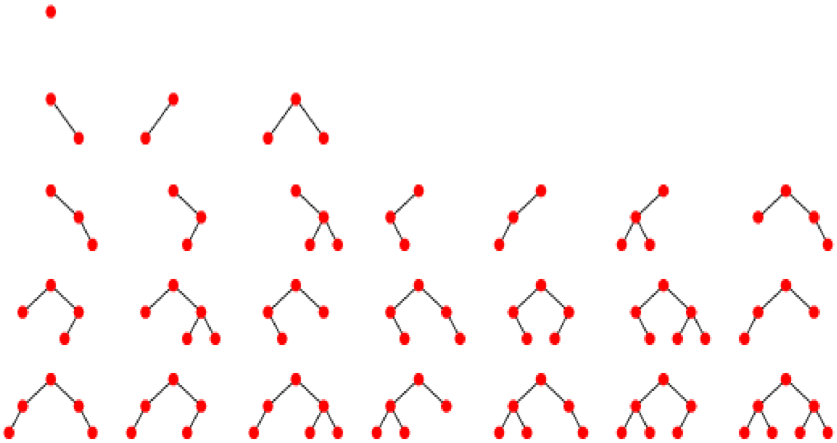


A tree (also called a bicentral tree) having two nodes that are graph centers. The numbers of bicentered trees on $n = 1,2, \ldots$ nodes are 0, 1, 0, 1, 1, 3, 4, 11, 20, 51, 108 ... (

● Binary Tree

A binary tree is a tree-like structure that is rooted and in which each vertex has at most two children and each child of a vertex is designated as its left or right child. In other words, unlike a proper tree, the relative position of the children is significant.

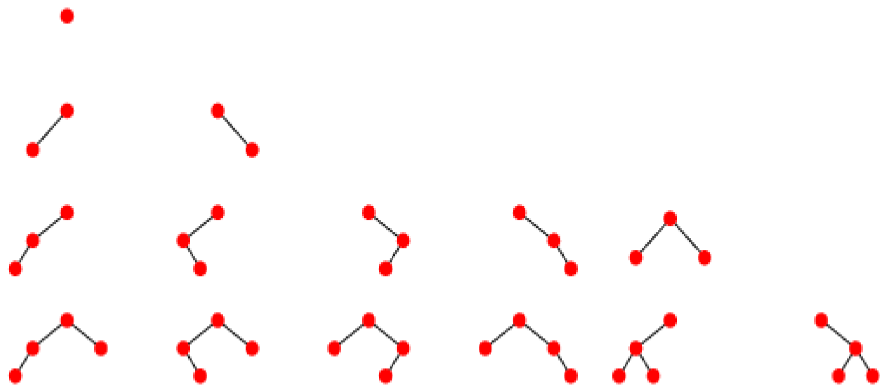
Dropping the requirement that left and right children are considered unique gives a true tree known as a weakly binary tree (in which, by convention, the root node is also required to be adjacent to at most one graph vertex).



The height of a binary tree is the number of levels within the tree. The numbers of binary trees of height $n = 1,2,\ldots$ nodes are 1, 3, 21, 651, 457653, ... A recurrence equation giving these counts is

$$a_n = a_{n-1}^2 + a_{n-1}\left(1 + \sqrt{4a_{n-1} - 3}\right)$$

with $a_1 = 1$.



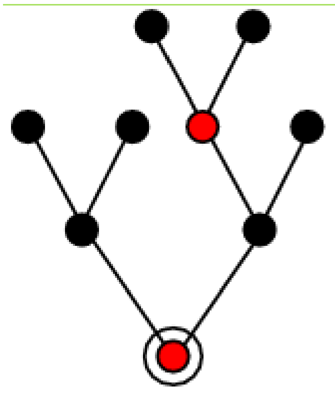
The number of binary trees with n nodes are $1, 2, 5, 14, 42, \ldots$ which are the Catalan number C_n.

For a binary tree of height h with n nodes,

$$h \leq n \leq 2^h - 1$$

These extremes correspond to a balanced tree (each node except the tree leaves has a left and right child, and all tree leaves are at the same level) and a degenerate tree (each node has only one outgoing branch), respectively.

● Red-Black Tree



An extended rooted binary tree satisfying the following conditions:

1. Every node has two children, each colored either red or black.

2. Every tree leaf node is colored black.

3. Every red node has both of its children colored black.

4. Every path from the root to a tree leaf contains the same number (the "black-height") of black nodes.

Let n be the number of internal nodes of a red-black tree. Then the number of red-black trees for $n = 1,2, \ldots$. is $2, 2, 3, 8, 14, 20, 35, 64, 122, \ldots$.

Let T_n be the generating function for the number of red-black trees of black-height h indexed by the number of tree leaves. Then $T_{h+1}(x) = [T_h(x)]^2 + [T_h(x)]^4$

Where $T_1(x) = x + x^2$.

If $T(x)$ is the generating function for the number of red-black trees, then $T(x) = x + x^2 + T(x^2(1 + x)^2)$

Let $rb(n)$ be the number of red-black trees with n tree leaves, $r(n)$ the number of red-rooted trees, and $b(n)$ the number of black-rooted trees. All three of the quantities satisfy the recurrence relation

$$R(n) = \sum_{n/4 \leq n \leq n/2} \binom{2m}{n-2m} R(m)$$

Where $\binom{n}{k}$ is a binomial coefficient,

$rb(1) = 1, rb(2) = 2$ for $R(n) = rb(n), r(1) = r(3) = 0, r(2) = 1$ for $R(n) = r(n)$, and $b(1) = 1$ for $R(n) = b(n)$.

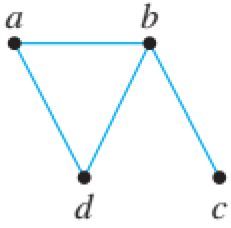
Exercise Set (9)

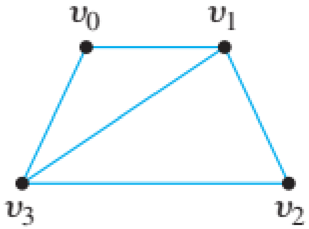
1. Prove that any tree is a bipartite graph.

2. Find the number of all spanning trees in

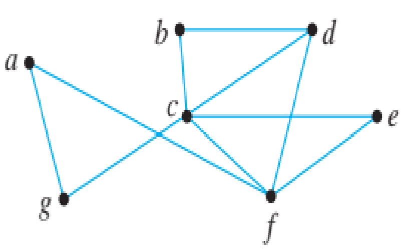
(a) K4; (b) $K_{2,3}$.

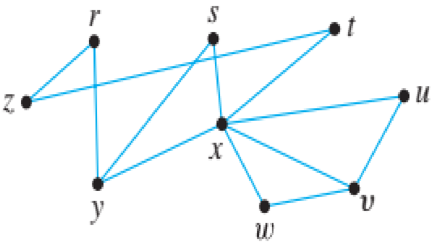
3. Find all possible spanning trees for each of the following graphs.





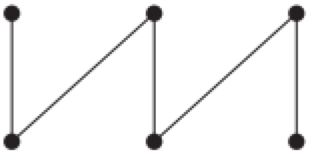
4. Find a spanning tree for each of the following graphs.



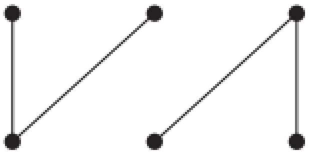


5. Which of these graphs are trees?

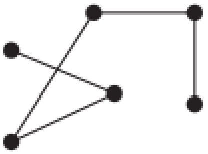
a)



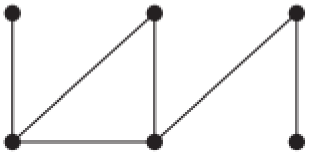
b)



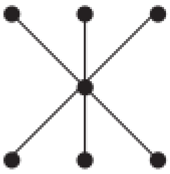
c)



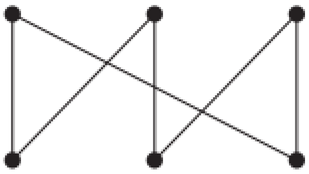
d)



e)

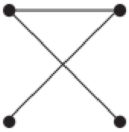


f)

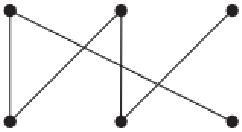


6. Which of these graphs are trees?

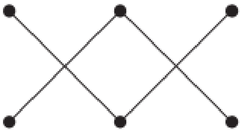
a)



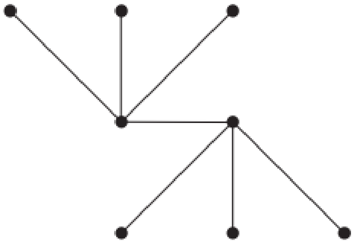
b)



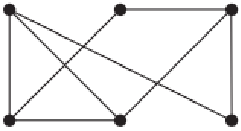
c)



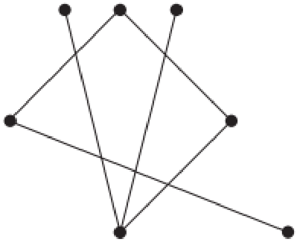
d)



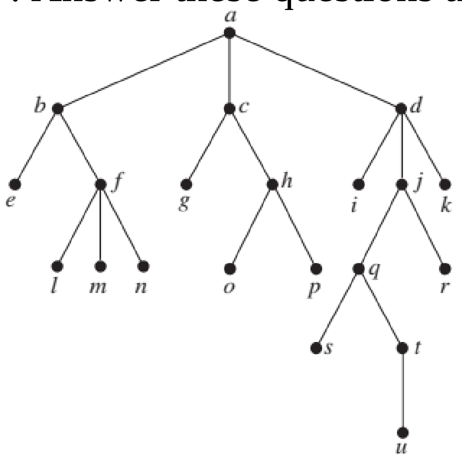
e)

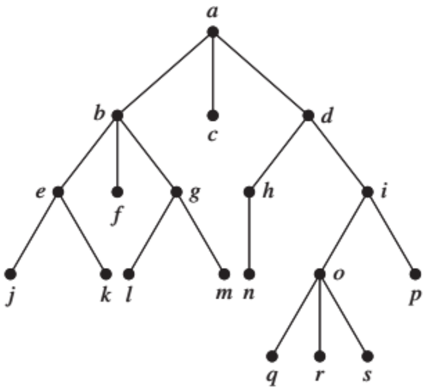


f)



7. Answer these questions about the rooted tree illustrated





a) Which vertex is the root?

b) Which vertices are internal?

c) Which vertices are leaves?

d) Which vertices are children of j?

e) Which vertex is the parent of h?

f) Which vertices are siblings of o?

g) Which vertices are ancestors of m?

h) Which vertices are descendants of b?

8. What is the level of each vertex of the rooted tree in Exercise 7?

9. Draw the subtree of the tree in Exercise 7 that is rooted at a) a. b) c. c) e.

10. How many non-isomorphic unrooted trees are there with three vertices?

11. How many non-isomorphic unrooted trees are there with four vertices?

12. How many edges does a tree with 10,000 vertices have?

13. How many vertices does a full 5-ary tree with 100 internal vertices have?

14. How many edges does a full binary tree with 1000 internal vertices have?

15. How many leaves does a full 3-ary tree with 100 vertices have?

16. How many edges are there in a forest of t trees containing a total of n vertices?

Bibliography

● V. K. Balakrishnan, *Theory and Problems of Combinatorics,* Schaum's Outline Series, McGraw-Hill, 1995.

● A. Engel, *Problem-Solving Strategies*, Springer Verlag, 1998.

● Susanna S. Epp , *Discrete Mathematics with Applications,* Fifth Edition, 2018 © Cengage-USA.

● M. Kac and S. M. Ulam, *Mathematics and Logic*, Dover Publications, NY, 1992.

● S. Lipschutz and M. L. Lipson, *2000 Solved Problems in Discrete Mathematics*, McGraw-Hill, 1992 .

● C.L. Liu, *Elements of Discrete Mathematics*, Second Edition, McGraw-Hill Book Company 1986.

●Kenneth H. Rosen, *Discrete mathematics and its applications*, 7th ed. Published by McGraw-Hill, a business unit of The McGraw-Hill Companies, Inc., 1221 Avenue of the Americas, New York, NY 10020. Copyright © 2012 by The McGraw-Hill Companies, Inc.

● A. Soifer, *Geometric Etudes in Combinatorial Mathematics,* Springer, 2010 (2nd, expanded edition).

●J. A. Bondy and U. S. R. Murty, *GRAPH THEORY WITH APPLICATIONS,* NORfH-HOLLAND New York • Amsterdam • Oxford.