Counting

Chapter 6

Mr. Shoaib Raza

Chapter Summary

- The Basics of Counting
- The Pigeonhole Principle
- Permutations and Combinations
- Binomial Coefficients and Identities
- Generalized Permutations and Combinations

The Basics of Counting

Section 6.1

COMBINATORICS

- Combinatorics is the mathematics of counting and arranging objects. Counting of objects with certain properties (enumeration) is required to solve many different types of problem.
- Applications, include topics as diverse as codes, circuit design and algorithm complexity [and gambling]

Counting

- Enumeration, the counting of objects with certain properties, is an important part of combinatorics.
- We must count objects to solve many different types of problems. For example, counting is used to:

1. Determine number of ordered or unordered arrangement of objects.
2. Generate all the arrangements of a specified kind which is important in computer simulations.
3. Compute probabilities of events.
4. Analyze the chance of winning games, lotteries etc.
5. Determine the complexity of algorithms.

Section Summary

- The Sum Rule
- The Product Rule
- The Subtraction Rule
- The Division Rule
- Examples, Examples, and Examples
- Tree Diagrams

Basic Counting Principles: The Sum Rule

The Sum Rule: If a task can be done either in one of n_1 ways or in one of n_2 ways to do the second task, where none of the set of n_1 ways is the same as any of the n_2 ways, then there are $n_1 + n_2$ ways to do the task.

The Sum Rule in terms of sets.

- The sum rule can be phrased in terms of sets.

  $|A \cup B| = |A| + |B|$ as long as A and B are disjoint sets.

- Or more generally,

$$|A_1 \cup A_2 \cup \cdots \cup A_m| = |A_1| + |A_2| + \cdots + |A_m|$$

when $A_i \cap A_j = \emptyset$ for all i, j.

- The case where the sets have elements in common will be discussed when we consider the subtraction rule and taken up fully in Chapter 8.

Basic Counting Principles: The Sum Rule

Example:

Suppose there are 7 different optional courses in Computer Science and 3 different optional courses in Mathematics. How many ways student can choose a course.

Solution: By the sum rule it follows that there are $7 + 3 = 10$ choices for a student who wants to take one optional course.

Basic Counting Principles: The Sum Rule

Example: The mathematics department must choose either a student or a faculty member as a representative for a university committee. How many 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: By the sum rule it follows that there are $37 + 83 = 120$ possible ways to pick a representative.

Basic Counting Principles: The Sum Rule

Example: A student can choose a computer project from one of the three lists. The three lists contain 23, 15 and 19 possible projects, respectively. How many possible projects are there to choose from?

Solution: The student can choose a project from the first list in 23 ways, from the second list in 15 ways, and from the third list in 19 ways. Hence, there are

$23 + 15 + 19 = 57$ projects to choose from.

Basic Counting Principles: The Product Rule

The Product Rule: A procedure can be broken down into a sequence of two tasks. There are n_1 ways to do the first task and n_2 ways to do the second task. Then there are $n_1 \cdot n_2$ ways to do the procedure.

Product Rule in Terms of Sets

- If $A_1, A_2, \ldots, A_m$ are finite sets, then the number of elements in the Cartesian product of these sets is the product of the number of elements of each set.
- The task of choosing an element in the Cartesian product $A_1 \times A_2 \times \cdots \times A_m$ is done by choosing an element in A_1, an element in A_2 , …, and an element in A_m.
- By the product rule, it follows that:
  $|A_1 \times A_2 \times \cdots \times A_m| = |A_1| \cdot |A_2| \cdot \cdots \cdot |A_m|$.

The Product Rule

Example: How many ways a student can choose one optional course each from computer science and mathematics courses if there are 7 different optional courses in Computer Science and 3 different optional courses in Mathematics.

Solution:

A student who wants to take one optional course of each subject, there are $7 \times 3 = 21$ choices.

The Product Rule

Example: The chairs of an auditorium are to be labeled with two characters, a letter followed by a digit. What is the largest number of chairs that can be labeled differently?

Solution:

The procedure of labeling a chair consists of two events, namely,

Assigning one of the 26 letters: A, B, C, …, Z and

Assigning one of the 10 digits: 0, 1, 2, …, 9

By product rule, there are $26 \times 10 = 260$ different ways that a chair can be labeled by both a letter and a digit.

The Product Rule

Example: Find the number n of ways that an organization consisting of 15 members can elect a president, treasurer, and secretary. (assuming no person is elected to more than one position)

Solution:

The president can be elected in 15 different ways; following this, the treasurer can be elected in 14 different ways; and following this, the secretary can be elected in 13 different ways. Thus, by product rule, there are

$$n = 15 \times 14 \times 13 = 2730$$

different ways in which the organization can elect the officers.

The Product Rule

Example: There are four bus lines between A and B; and three bus lines between B and C.

Find the number of ways a person can travel:

a) By bus from A to C by way of B;

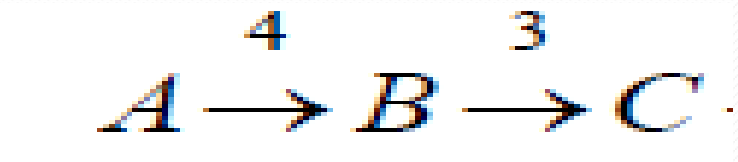
b) Round trip by bus from A to C by way of B;

c) Round trip by bus from A to C by way of B, if the person does not want to use a bus line more than once.

The Product Rule

a) By bus from A to C by way of B;

Solution:



There are 4 ways to go from A to B and 3 ways to go from B to C; hence there are $4 \times 3 = 12$ ways to go from A to C by way of B.

The Product Rule

b) Round trip by bus from A to C by way of B;

Solution:

The person will travel from A to B to C to B to A for the round trip. i.e. (A →B →C →B →A)

The person can travel 4 ways from A to B and 3 way from B to C and back.

Thus there are 4 ×3 ×3 × 4 = 144 ways to travel the round trip.

The Product Rule

c) Round trip by bus from A to C by way of B, if the person does not want to use a bus line more than once.

Solution:

$$A \xrightarrow{4} B \xrightarrow{3} C \xrightarrow{2} B \xrightarrow{3} A$$

The person can travel 4 ways from A to B and 3 ways from B to C, but only 2 ways from C to B and 3 ways from B to A, since bus line cannot be used more than once. Hence there are

$4 \times 3 \times 2 \times 3 = 72$ ways

to travel the round trip without using a bus line more than once.

The Product Rule

Example: A bit string is a sequence of 0's and 1's.How many bit string are there of length 4?

Solution:

Each bit (binary digit) is either 0 or 1.

Hence, there are **2** ways to choose each bit. Since we have to choose four bits therefore,

$$2 \times 2 \times 2 \times 2 = 2^4 = 16$$

the product rule shows, there are a total of different bit strings of length four.

The Product Rule

Example: How many bit strings of length 8:

(i) begin with a 1?

(ii) begin and end with a 1?

Solution:

(i) If the first bit (left most bit) is a 1, then it can be filled in only one way. Each of the remaining seven positions in the bit string can be filled in 2 ways (i.e., either by 0 or 1). Hence, there are

$$1 \times 2 \times 2 \times 2 \times 2 \times 2 \times 2 \times 2 = 2^7 = 128$$

different bit strings of length 8 that begin with a 1.

The Product Rule

(ii) begin and end with a 1?

Solution:

If the first and last bit in an 8 bit string is a 1, then only the intermediate six bits can be filled in 2 ways, i.e. by a 0 or 1. Hence there are

$$1 \times 2 \times 2 \times 2 \times 2 \times 2 \times 2 \times 1 = 2^6 = 64$$

different bit strings of length 8 that begin and end with a 1.

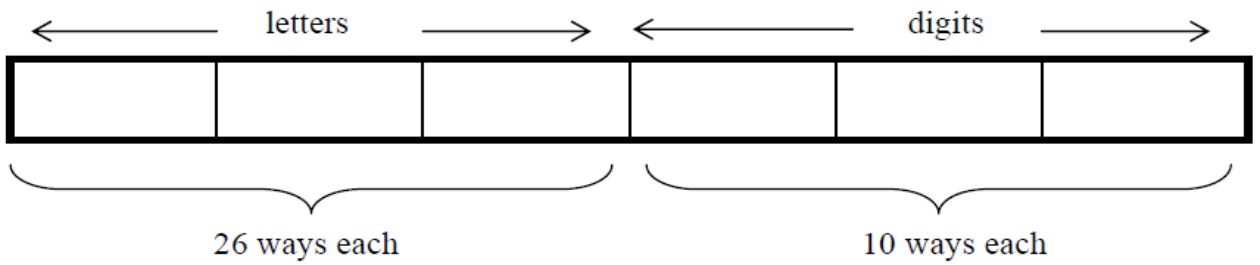
The Product Rule

Example: Suppose that an automobile license plate has three letters followed by three digits.

(a) How many different license plates are possible?

Solution:

Each of the three letters can be written in 26 different ways, and each of the three digits can be written in 10 different ways.



Hence, by the product rule, there is a total of

$26 \times 26 \times 26 \times 10 \times 10 \times 10 = 17{,}576{,}000$

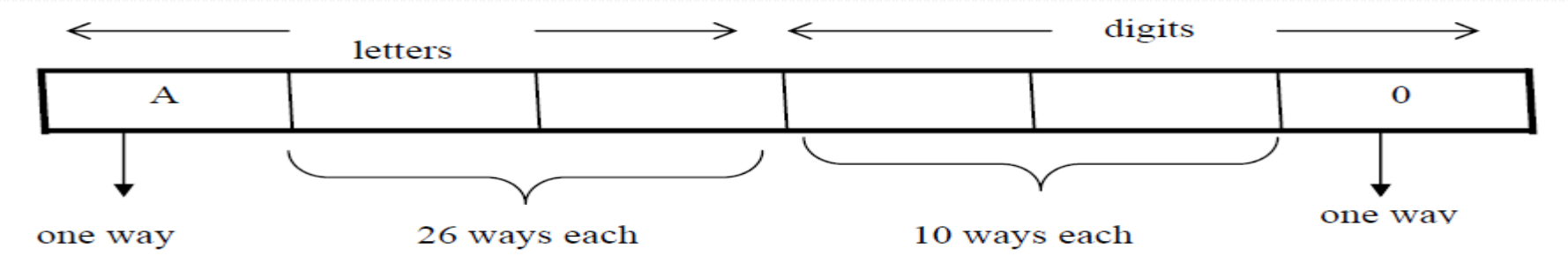
different License plates possible.

The Product Rule

(b) How many license plates could begin with A and end on 0?

Solution:

The first and last place can be filled in one way only, while each of second and third place can be filled in 26 ways and each of fourth and fifth place can be filled in 10 ways.



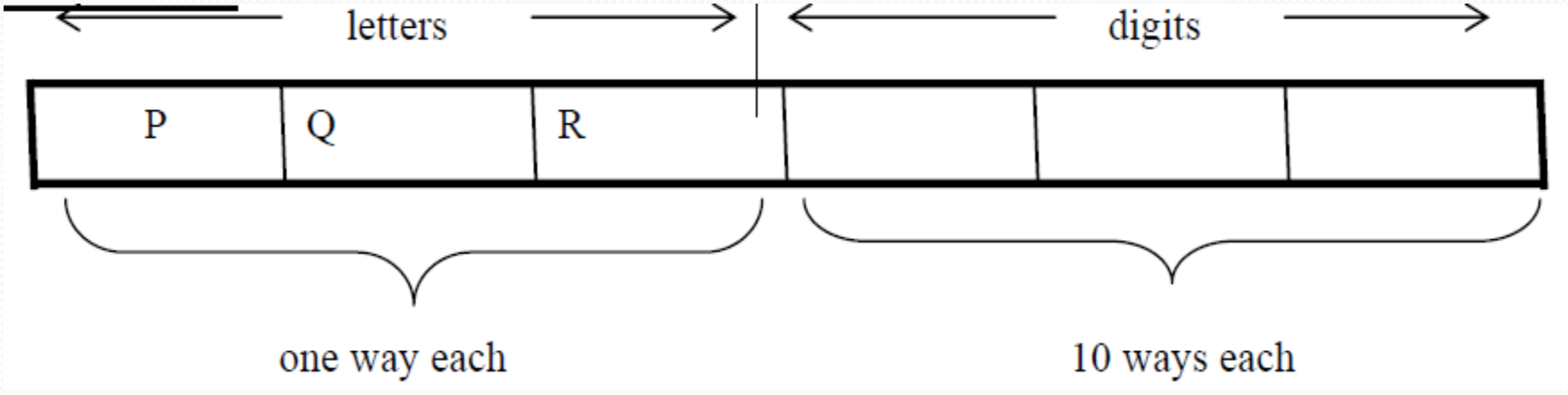
Number of license plates that begin with A and end in 0 are

$$1 \times 26 \times 26 \times 10 \times 10 \times 1 = 67600$$

The Product Rule

(c) How many license plates begin with PQR.

Solution:



Number of license plates that begin with PQR are

$$1 \times 1 \times 1 \times 10 \times 10 \times 10 = 1000 \text{ ways.}$$

The Product Rule

(d) How many license plates are possible in which all the letters and digits are distinct?

Solution:

The first letter place can be filled in 26 ways. Since, the second letter place should contain a different letter than the first, so it can be filled in 25 ways. Similarly, the third letter place can be filled in 24 ways. And the digits can be respectively filled in 10, 9, and 8 ways.

Hence;

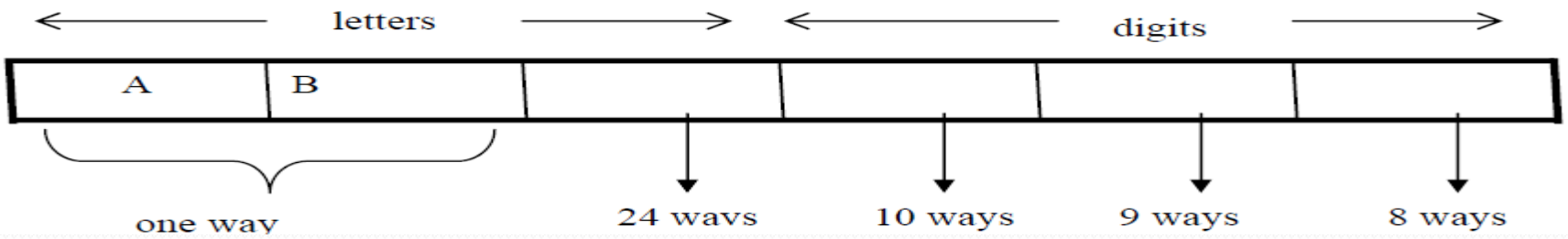
number of license plates in which all the letters and digits are distinct are

$$26 \times 25 \times 24 \times 10 \times 9 \times 8 = 11, 232, 000$$

The Product Rule

(e) How many license plates could begin with AB and have all three letters and digits distinct.

Solution:



The first two letters places are fixed (to be filled with A and B), so there is only one way to fill them. The third letter place should contain a letter different from A & B, so there

are 24 ways to fill it.

The three digit positions can be filled in 10 and 8 ways to have distinct digits. Hence, desired number of license plates are

$$1 \times 1 \times 24 \times 10 \times 9 \times 8 = 17280$$

Telephone Numbering Plan

Example: The *North American numbering plan* (*NANP*) specifies that a telephone number consists of 10 digits, consisting of a three-digit area code, a three-digit office code, and a four-digit station code. There are some restrictions on the digits.

- Let *X* denote a digit from 0 through 9.
- Let *N* denote a digit from 2 through 9.
- Let *Y* denote a digit that is 0 or 1.
- In the old plan (in use in the 1960s) the format was *NYX-NNX-XXXX*.
- In the new plan, the format is *NXX-NXX-XXXX*.

How many different telephone numbers are possible under the old plan and the new plan?

Solution: Use the Product Rule.

- There are $8 \cdot 2 \cdot 10 = 160$ area codes with the format *NYX*.
- There are $8 \cdot 10 \cdot 10 = 800$ area codes with the format *NXX*.
- There are $8 \cdot 8 \cdot 10 = 640$ office codes with the format *NNX*.
- There are $10 \cdot 10 \cdot 10 \cdot 10 = 10{,}000$ station codes with the format *XXXX*.

Number of old plan telephone numbers: $160 \cdot 640 \cdot 10{,}000 = 1{,}024{,}000{,}000$.

Number of new plan telephone numbers: $800 \cdot 800 \cdot 10{,}000 = 6{,}400{,}000{,}000$.

NUMBER OF ITERATIONS OF A NESTED LOOP

Example: Determine how many times the inner loop will be iterated when the following algorithm is implemented and run

For i: = 1 to 4

For j : = 1 to 3

[Statement in body of inner loop. None contain branching statements that lead out of the inner loop.]

next j

next i

Solution:

The outer loop is iterated four times, and during each iteration of the outer loop, there are three iterations of the inner loop.

Hence, by product rules the total number of iterations of inner loop is 4·3=12

Example: Determine how many times the inner loop will be iterated when the following algorithm is implemented and run.

```
for     i = 5 to 50
for     j: = 10 to 20
```

[Statement in body of inner loop. None contain branching statements that lead out of the inner loop.]

```
        next j
        next i
```

Solution:

The outer loop is iterated 50 - 5 + 1 = 46 times and during each iteration of the outer loop there are 20 - 10 + 1 = 11 iterations of the inner loop. Hence by product rule, the total number of iterations of the inner loop is 46×11 = 506.

Example: Determine how many times the inner loop will be iterated when the following algorithm is implemented and run.

```
for     i: = 1 to 4
for     j: = 1 to i
[Statements in body of inner loop. None contain branching statements that lead outside the loop.]
next j
next i
```

Solution:

The outer loop is iterated 4 times, but during each iteration of the outer loop, the inner loop iterates different number of times.

For first iteration of outer loop, inner loop iterates 1 times.

For second iteration of outer loop, inner loop iterates 2 times.

For third iteration of outer loop, inner loop iterates 3 times.

For fourth iteration of outer loop, inner loop iterates 4 times.

Hence, total number of iterations of inner loop = 1 + 2 + 3 + 4 = 10.

Combining the Sum and Product Rule

Example: Suppose statement labels in a programming language can be either a single letter or a letter followed by a digit. Find the number of possible labels.

Solution:

- First consider variable names one character in length. Since such names consist of a single letter, there are 26 variable names of length 1.
- Next, consider variable names two characters in length. Since the first character is a letter, there are 26 ways to choose it. The second character is a digit, there are 10 ways to choose it. Hence, to construct variable name of two characters in length, there are 26×10 = 260 ways.
- Finally, by sum rule, there are 26 + 260 = 286 possible variable names in the programming language.

Combining the Sum and Product Rule

Example: A computer access code word consists of from one to three letters of English alphabets with repetitions allowed. How many different code words are possible.

Solution:

Number of code words of length 1 = 26^1

Number of code words of length 2 = 26^2

Number of code words of length 3 = 26^3

Hence, the total number of code words =

$$26^1 + 26^2 + 26^3 = 18{,}278$$

Counting Passwords

- Combining the sum and product rule allows us to solve more complex problems.

Example: Each user on a computer system has a password, which is six to eight characters long, where each character is an uppercase letter or a digit. Each password must contain at least one digit. How many possible passwords are there?

Solution: Let P be the total number of passwords, and let P_6, P_7, and P_8 be the passwords of length 6, 7, and 8.

- By the sum rule $P = P_6 + P_7 + P_8$.

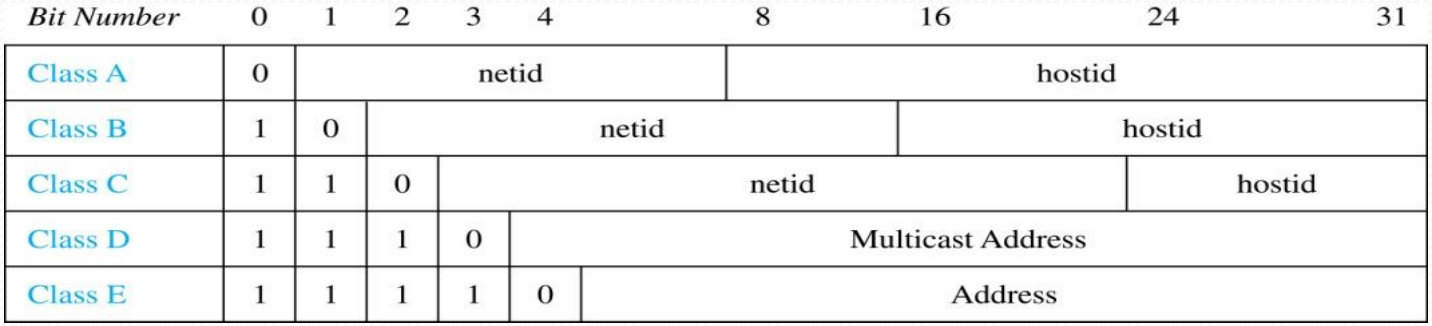
Finding P_6 directly is difficult. To find P_6 it is easier to find the number of strings of uppercase letters and digits that are six characters long, including those with no digits, and subtract from this the number of strings with no digits. By the product rule, the number of strings of six characters is 36^6 ,and the number of strings with no digits is 26^6

Counting Passwords(Continued)

- To find each of P6, P7, and P8 , we find the number of passwords of the specified length composed of letters and digits and subtract the number composed only of letters. We find that:
- $P_6 = 36^6 - 26^6 = 2{,}176{,}782{,}336 - 308{,}915{,}776 = 1{,}867{,}866{,}560.$
- $P_7 = 36^7 - 26^7 = 78{,}364{,}164{,}096 - 8{,}031{,}810{,}176 = 70{,}332{,}353{,}920.$
- $P_8 = 36^8 - 26^8 = 2{,}821{,}109{,}907{,}456 - 208{,}827{,}064{,}576 = 2{,}612{,}282{,}842{,}880.$
- Consequently, $P = P_6 + P_7 + P_8 = 2{,}684{,}483{,}063{,}360.$

Internet Addresses

- Version 4 of the Internet Protocol (IPv4) uses 32 bits.



Bit Number	0	1	2	3	4	8	16	24	31
Class A	0	netid				hostid			
Class B	1	0	netid				hostid		
Class C	1	1	0	netid				hostid	
Class D	1	1	1	0	Multicast Address				
Class E	1	1	1	1	0	Address			

- **Class A Addresses**: used for the largest networks, a 0,followed by a 7-bit netid and a 24-bit hostid.
- **Class B Addresses**: used for the medium-sized networks, a 10,followed by a 14-bit netid and a 16-bit hostid.
- **Class C Addresses**: used for the smallest networks, a 110,followed by a 21-bit netid and a 8-bit hostid.
  - Neither Class D nor Class E addresses are assigned as the address of a computer on the internet. Only Classes A, B, and C are available.
  - 1111111 is not available as the netid of a Class A network.
  - Hostids consisting of all 0s and all 1s are not available in any network.

Counting Internet Addresses

Example: How many different IPv4 addresses are available for computers on the internet?

Solution: Use both the sum and the product rule. Let x be the number of available addresses, and let x_A, x_B, and x_C denote the number of addresses for the respective classes.

- To find, x_A: $2^7 - 1 = 127$ netids. $2^{24} - 2 = 16{,}777{,}214$ hostids.
  $x_A = 127 \cdot 16{,}777{,}214 = 2{,}130{,}706{,}178.$
- To find, x_B: $2^{14} = 16{,}384$ netids. $2^{16} - 2 = 16{,}534$ hostids.
  $x_B = 16{,}384 \cdot 16{,}534 = 1{,}073{,}709{,}056.$
- To find, x_C: $2^{21} = 2{,}097{,}152$ netids. $2^8 - 2 = 254$ hostids.
  $x_C = 2{,}097{,}152 \cdot 254 = 532{,}676{,}608.$
- Hence, the total number of available IPv4 addresses is

$$\begin{aligned} x &= x_A + x_B + x_C \\ &= 2{,}130{,}706{,}178 + 1{,}073{,}709{,}056 + 532{,}676{,}608 \\ &= 3{,}737{,}091{,}842. \end{aligned}$$

Not Enough Today !!
The newer IPv6 protocol solves the problem of too few addresses.

Basic Counting Principles: Subtraction Rule

Subtraction Rule: If a task can be done either in one of n_1 ways or in one of n_2 ways, then the total 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.

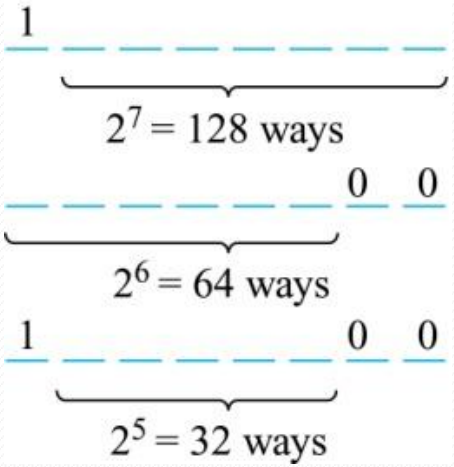
- Also known as, the *principle of inclusion-exclusion*:

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

Counting Bit Strings

Example: How many bit strings of length eight start either with a 1 bit or end with the two bits 00?

Solution: Use the subtraction rule.



- Number of bit strings of length eight that start with a 1 bit: $2^7 = 128$
- Number of bit strings of length eight that end with bits 00: $2^6 = 64$
- Number of bit strings of length eight that start with a 1 bit and end with bits 00 : $2^5 = 32$

Hence, the number is $128 + 64 - 32 = 160$.

Counting Functions

Counting Functions: How many functions are there from a set with m elements to a set with n elements?

Solution: Since a function represents a choice of one of the n elements of the codomain for each of the m elements in the domain, the product rule tells us that there are $n \cdot n \cdots n = n^m$ such functions.

Counting One-to-One Functions: How many one-to-one functions are there from a set with m elements to one with n elements?

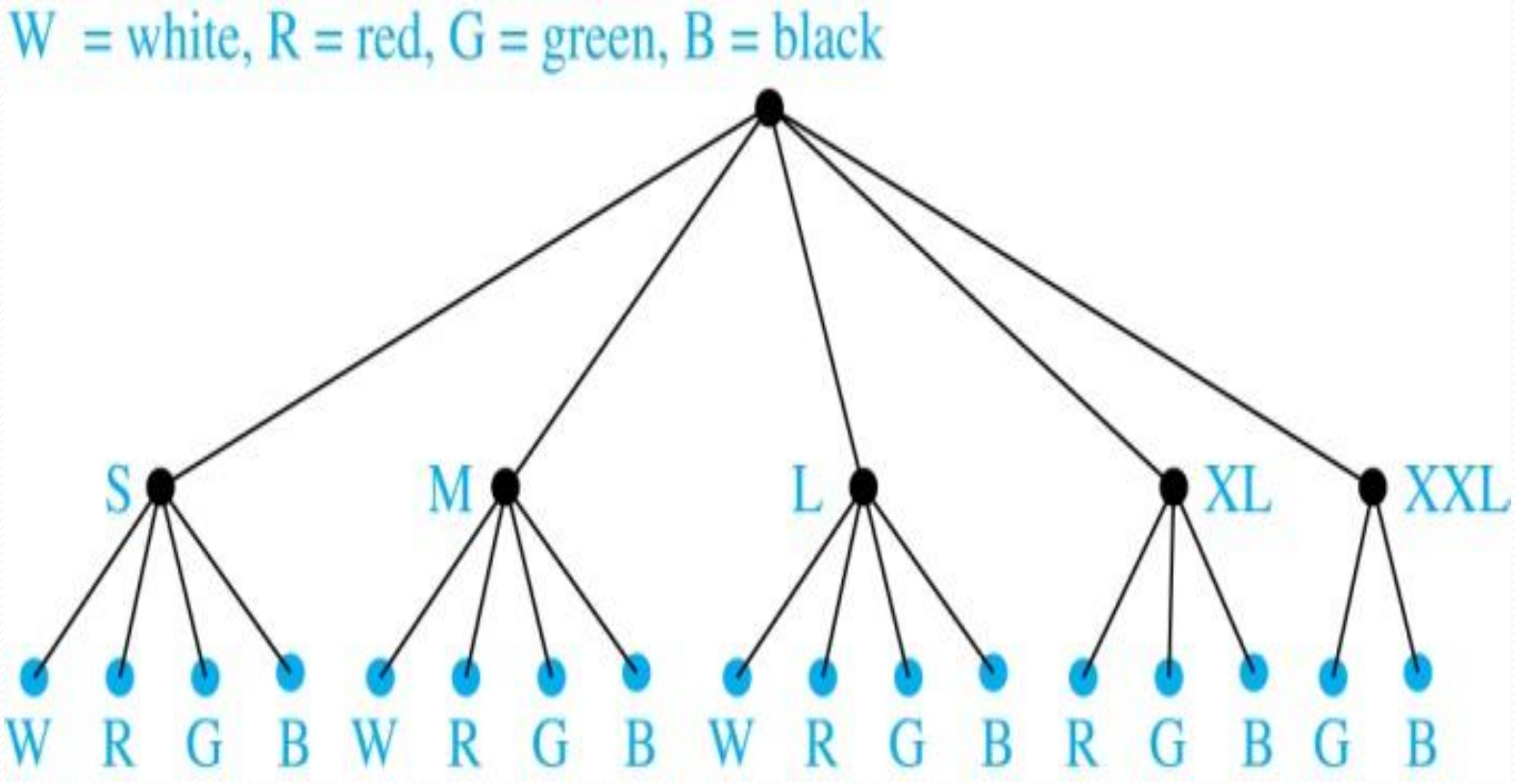
Solution: Suppose the elements in the domain are $a_1, a_2, \ldots, a_m$. There are n ways to choose the value of a_1 and $n-1$ ways to choose a_2, etc. The product rule tells us that there are $n(n-1)(n-2)\cdots(n-m+1)$ such functions.

Tree Diagrams

- **Tree Diagrams**: We can solve many counting problems through the use of *tree diagrams*, where a branch represents a possible choice and the leaves represent possible outcomes.
- **Example**: Suppose that "I Love Discrete Math" T-shirts come in five different sizes: S,M,L,XL, and XXL. Each size comes in four colors (white, red, green, and black), except XL, which comes only in red, green, and black, and XXL, which comes only in green and black. What is the minimum number of shirts that the campus book store needs to stock to have one of each size and color available?

Tree Diagrams

- **Solution**: Draw the tree diagram.



- The store must stock 17 T-shirts.

The Pigeonhole Principle

Section 6.2

Section Summary

- The Pigeonhole Principle
- The Generalized Pigeonhole Principle

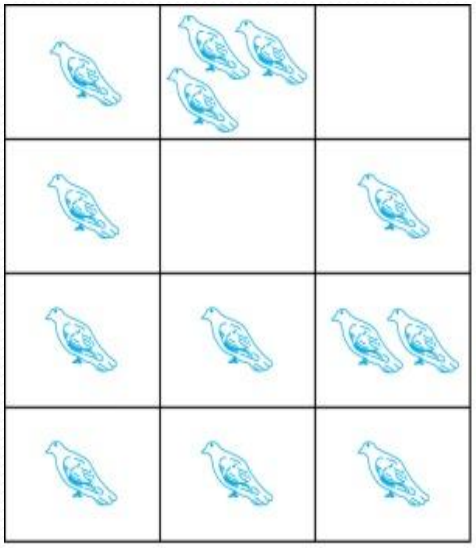
The Pigeonhole Principle

Pigeonhole Principle: If k is a positive integer and $k + 1$ objects are placed into k boxes, then at least one box contains two or more objects.

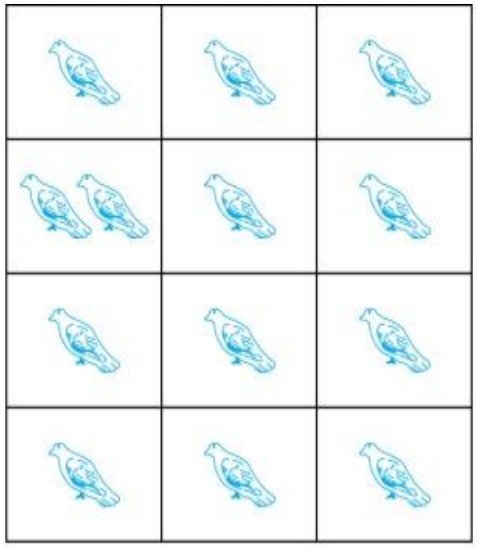
Proof: We use a proof by contraposition. Suppose none of the k boxes has more than one object. Then the total number of objects would be at most k. This contradicts the statement that we have $k + 1$ objects.

The Pigeonhole Principle

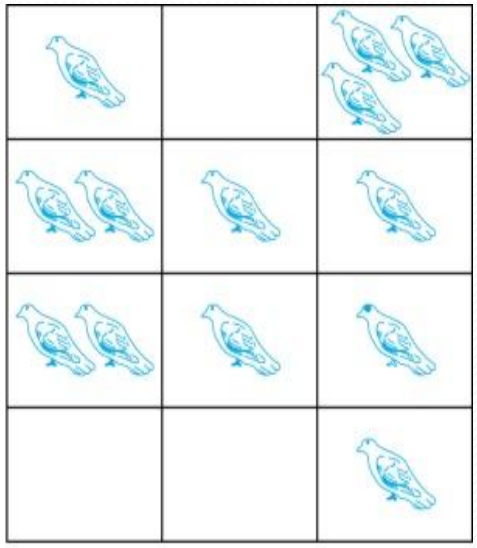
- If a flock of 20 pigeons roosts in a set of 19 pigeonholes, one of the pigeonholes must have more than 1 pigeon.



(a)



(b)



(c)

The Pigeonhole Principle

Corollary 1: A function f from a set with $k + 1$ elements to a set with k elements is not one-to-one.

Proof: Use the pigeonhole principle.

- Create a box for each element y in the codomain of f.
- Put in the box for y all of the elements x from the domain such that $f(x) = y$.
- Because there are $k + 1$ elements and only k boxes, at least one box has two or more elements.

Hence, f can't be one-to-one. ◀

The Generalized Pigeonhole Principle

The Generalized Pigeonhole Principle: If N objects are placed into k boxes, then there is at least one box containing at least $\lceil N/k \rceil$ objects.

Proof: We use a proof by contraposition. Suppose that none of the boxes contains more than $\lceil N/k \rceil - 1$ objects. Then the total number of objects is at most

$$k\left(\left\lceil \frac{N}{k} \right\rceil - 1\right) < k\left(\left(\frac{N}{k} + 1\right) - 1\right) = N,$$

where the inequality $\lceil N/k \rceil < \lceil N/k \rceil + 1$ has been used. This is a contradiction because there are a total of n objects. ◀

Pigeonhole Principle

Example: Among any group of 367 people, there must be at least two with the same birthday, because there are only 366 possible birthdays. $\lceil 367/366 \rceil = 2$

Example: Among 100 people there are at least $\lceil 100/12 \rceil = 9$ who were born in the same month.

Example: In any set of 27 English , must be at least two that begin with the same letter, since there are 26 letters in the English alphabet. $\lceil 27/26 \rceil = 2$

The Generalized Pigeonhole Principle

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 guarantee that at least six students receive the same grade is the smallest integer N such that $\lceil N/K \rceil = \lceil N/5 \rceil = 6$. The smallest such integer is

$$N = K(\lceil N/K \rceil - 1) + 1 = 5(6-1)+1 = 5 \cdot 5 + 1 = 26.$$

Thus 26 is the minimum number of students needed to be sure that at least 6 students will receive the same grades.

Permutations and Combinations

Section 6.3

Section Summary

- Permutations
- Combinations
- Combinatorial Proofs

Permutations

Definition: A *permutation* of a set of distinct objects is an ordered arrangement of these objects. An ordered arrangement 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.

- The number of r-permutations of a set with n elements is denoted by $P(n,r)$.
- The 2-permutations of $S = \{1,2,3\}$ are 1,2; 1,3; 2,1; 2,3; 3,1; and 3,2. Hence, $P(3,2) = 6$.

A Formula for the Number of Permutations

Theorem 1: If n is a positive integer and r is an integer with $1 \leq r \leq n$, then there are

$$P(n, r) = n(n - 1)(n - 2) \cdots (n - r + 1)$$

r-permutations of a set with n distinct elements.

Proof: Use the product rule. The first element can be chosen in n ways. The second in $n - 1$ ways, and so on until there are $(n - (r - 1))$ ways to choose the last element.

- Note that $P(n,0) = 1$, since there is only one way to order zero elements.

Corollary 1: If n and r are integers with $1 \leq r \leq n$, then

$$P(n, r) = \frac{n!}{(n-r)!}$$

Solving Counting Problems by Counting Permutations

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:

$$P(100,3) = 100 \cdot 99 \cdot 98 = 970{,}200$$

Solving Counting Problems by Counting Permutations (*continued*)

- **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.

Solving Counting Problems by Counting Permutations (*continued*)

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 first city is chosen, and the rest are ordered arbitrarily. Hence the orders are:

$$P(7,7) = 7! = 7 \cdot 6 \cdot 5 \cdot 4 \cdot 3 \cdot 2 \cdot 1 = 5040$$

If she wants to find the tour with the shortest path that visits all the cities, she must consider 5040 paths!

Solving Counting Problems by Counting Permutations (*continued*)

Example: How many permutations of the letters *ABCDEFGH* contain the string *ABC* ?

Solution: We solve this problem by counting the permutations of six objects, *ABC*, *D*, *E*, *F*, *G*, and *H*.

$P(6,6) = 6! = 6 \cdot 5 \cdot 4 \cdot 3 \cdot 2 \cdot 1 = 720$

Combinations

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.

- The number of *r*-combinations of a set with n distinct elements is denoted by $C(n, r)$.

- The notation $\binom{n}{r}$ is also used and is called a *binomial coefficient*. (*We will see the notation again in the binomial theorem in Section*. 6.4)

Combinations

Example:

- Let S be the set $\{a, b, c, d\}$. Then $\{a, c, d\}$ is a 3-combination from S. It is the same as $\{d, c, a\}$ since the order listed does not matter.
- $C(4,2) = 6$ because the 2-combinations of $\{a, b, c, d\}$ are the six subsets $\{a, b\}$, $\{a, c\}$, $\{a, d\}$, $\{b, c\}$, $\{b, d\}$, and $\{c, d\}$.

Combinations

Theorem 2: The number of r-combinations of a set with n elements, where $n \geq r \geq 0$, equals

$$C(n,r) = \frac{n!}{(n-r)!r!}.$$

Proof: By the product rule $P(n, r) = C(n,r) \cdot P(r,r)$. Therefore,

$$C(n,r) = \frac{P(n,r)}{P(r,r)} = \frac{n!/(n-r)!}{r!/(r-r)!} = \frac{n!}{(n-r)!r!}.$$

Combinations

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 deck of 52 cards?

Solution: Since the order in which the cards are dealt does not matter, the number of five card hands is:

$$C(52,5) = \frac{52!}{5!47!}$$
$$= \frac{52\cdot 51\cdot 50\cdot 49\cdot 48}{5\cdot 4\cdot 3\cdot 2\cdot 1} = 26\cdot 17\cdot 10\cdot 49\cdot 12 = 2,598,960$$

- The different ways to select 47 cards from 52 is

$$C(52,47) = \frac{52!}{47!5!} = C(52,5) = 2,598,960.$$

This is a special case of a general result. →

Combinations

Corollary 2: Let n and r be nonnegative integers with $r \leq n$. Then $C(n, r) = C(n, n - r)$.

Proof: From Theorem 2, it follows that

$$C(n, r) = \frac{n!}{(n-r)!r!}$$

and

$$C(n, n - r) = \frac{n!}{(n-r)![n-(n-r)]!} = \frac{n!}{(n-r)!r!} .$$

Hence, $C(n, r) = C(n, n - r)$. ◀

This result can be proved without using algebraic manipulation. →

Combinatorial Proofs

- **Definition 1**: A *combinatorial proof* of an identity is a proof that uses one of the following methods.
  - A *double counting proof* uses counting arguments to prove that both sides of an identity count the same objects, but in different ways.
  - A *bijective proof* shows that there is a bijection between the sets of objects counted by the two sides of the identity.

Combinatorial Proofs

- Here are two combinatorial proofs that

$$C(n, r) = C(n, n - r)$$

when r and n are nonnegative integers with $r < n$:

  - *Bijective Proof*: Suppose that S is a set with n elements. The function that maps a subset A of S to $\bar{A}$ is a bijection between the subsets of S with r elements and the subsets with $n - r$ elements. Since there is a bijection between the two sets, they must have the same number of elements. ◀
  - *Double Counting Proof*: By definition the number of subsets of S with r elements is $C(n, r)$. Each subset A of S can also be described by specifying which elements are not in A, i.e., those which are in $\bar{A}$. Since the complement of a subset of S with r elements has $n - r$ elements, there are also $C(n, n - r)$ subsets of S with r elements. ◀

Combinations

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: By Theorem 2, the number of combinations is

$$C(10,5) = \frac{10!}{5!5!} = 252.$$

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?

Solution: By Theorem 2, the number of possible crews is

$$C(30,6) = \frac{30!}{6!24!} = \frac{30\cdot 29\cdot 28\cdot 27\cdot 26\cdot 25}{6\cdot 5\cdot 4\cdot 3\cdot 2\cdot 1} = 593,775\ .$$

Binomial Coefficients and Identities

Section 6.4

Section Summary

- The Binomial Theorem
- Pascal's Identity and Triangle

Binomial Theorem

Binomial Theorem: Let x and y be variables, and n a nonnegative integer. Then:

$$(x+y)^n = \sum_{j=0}^{n} \binom{n}{j} x^{n-j}y^j = \binom{n}{0}x^n + \binom{n}{1}x^{n-1}y + \cdots + \binom{n}{n-1}xy^{n-1} + \binom{n}{n}y^n.$$

Proof: We use combinatorial reasoning . The terms in the expansion of $(x + y)^n$ are of the form $x^{n-j}y^j$ for $j = 0,1,2,\ldots,n$. To form the term $x^{n-j}y^j$, it is necessary to choose $n-j$ xs from the n sums. Therefore, the coefficient of $x^{n-j}y^j$ is $\binom{n}{n-j}$ which equals $\binom{n}{j}$. ◄

Powers of Binomial Expressions

Definition: A *binomial* expression is the sum of two terms, such as $x + y$. (More generally, these terms can be products of constants and variables.)

- We can use counting principles to find the coefficients in the expansion of $(x + y)^n$ where n is a positive integer.
- To illustrate this idea, we first look at the process of expanding $(x + y)^3$.
- $(x + y)\ (x + y)\ (x + y)$ expands into a sum of terms that are the product of a term from each of the three sums.
- Terms of the form x^3, x^2y, $x\,y^2$, y^3 arise. The question is what are the coefficients?
  - To obtain x^3, an x must be chosen from each of the sums. There is only one way to do this. So, the coefficient of x^3 is 1.
  - To obtain x^2y, an x must be chosen from two of the sums and a y from the other. There are $\binom{3}{2}$ ways to do this and so the coefficient of x^2y is 3.
  - To obtain xy^2, an x must be chosen from of the sums and a y from the other two . There are $\binom{3}{1}$ ways to do this and so the coefficient of xy^2 is 3.
  - To obtain y^3, a y must be chosen from each of the sums. There is only one way to do this. So, the coefficient of y^3 is 1.
- We have used a counting argument to show that $(x + y)^3 = x^3 + 3x^2y + 3x\,y^2 + y^3$.
- Next we present the binomial theorem gives the coefficients of the terms in the expansion of $(x + y)^n$.

Using the Binomial Theorem

Example:

What is the expansion of $(x + y)^4$?

Solution: From the binomial theorem it follows that

$$(x+y)^4 = \sum_{j=0}^{4} \binom{4}{j} x^{4-j} y^j$$

$$= \binom{4}{0} x^4 + \binom{4}{1} x^3 y + \binom{4}{2} x^2 y^2 + \binom{4}{3} x y^3 + \binom{4}{4} y^4$$

$$= x^4 + 4x^3 y + 6x^2 y^2 + 4xy^3 + y^4.$$

Using the Binomial Theorem

What is the coefficient of $x^{12}y^{13}$ in the expansion of $(x+y)^{25}$?

Solution: From the binomial theorem it follows that this coefficient is

$$\binom{25}{13} = \frac{25!}{13!\,12!} = 5{,}200{,}300.$$

Using the Binomial Theorem

Example: What is the coefficient of $x^{12}y^{13}$ in the expansion of $(2x - 3y)^{25}$?

Solution: We view the expression as $(2x + (-3y))^{25}$. By the binomial theorem

$$(2x + (-3y))^{25} = \sum_{j=0}^{25} \binom{25}{j} (2x)^{25-j}(-3y)^{j}.$$

Consequently, the coefficient of $x^{12}y^{13}$ in the expansion is obtained when $j = 13$.

$$\binom{25}{13} 2^{12}(-3)^{13} = -\frac{25!}{13!12!}2^{12}3^{13}.$$

A Useful Identity

Corollary 1: With $n \geq 0$, $\sum_{k=0}^{n} \binom{n}{k} = 2^n$.

Proof (*using binomial theorem*): With $x = 1$ and $y = 1$, from the binomial theorem we see that:

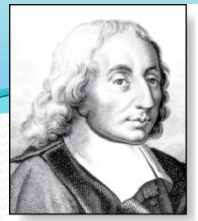
$$2^n = (1+1)^n = \sum_{k=0}^{n} \binom{n}{k} 1^k 1^{(n-k)} = \sum_{k=0}^{n} \binom{n}{k}. \quad \blacktriangleleft$$

Proof (*combinatorial*): Consider the subsets of a set with n elements. There are $\binom{n}{0}$ subsets with zero elements, $\binom{n}{1}$ with one element, $\binom{n}{2}$ with two elements, …, and $\binom{n}{n}$ with n elements. Therefore the total is $\sum_{k=0}^{n} \binom{n}{k}$.

Since, we know that a set with n elements has 2^n subsets, we conclude:

$$\sum_{k=0}^{n} \binom{n}{k} = 2^n. \quad \blacktriangleleft$$

Blaise Pascal
(1623-1662)



Pascal's Identity

Pascal's Identity: If n and k are integers with $n \geq k \geq 0$, then

$$\binom{n+1}{k} = \binom{n}{k-1} + \binom{n}{k}.$$

Proof (*combinatorial*): Let T be a set where $|T| = n + 1$, $a \in T$, and $S = T - \{a\}$. There are $\binom{n+1}{k}$ subsets of T containing k elements. Each of these subsets either:

- contains a with $k - 1$ other elements, or
- contains k elements of S and not a.

There are

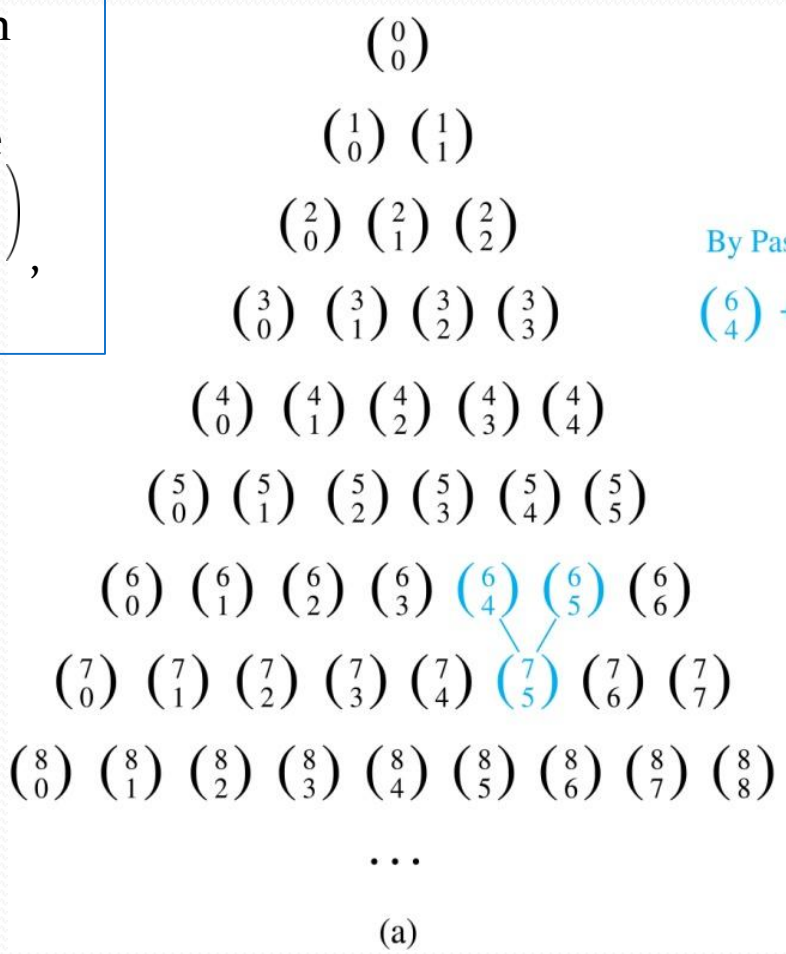
- $\binom{n}{k-1}$ subsets of k elements that contain a, since there are $\binom{n}{k-1}$ subsets of $k - 1$ elements of S,
- $\binom{n}{k}$ subsets of k elements of T that do not contain a, because there are $\binom{n}{k}$ subsets of k elements of S.

Hence, $$\binom{n+1}{k} = \binom{n}{k-1} + \binom{n}{k}.$$ ◀

See Exercise 19 *for an algebraic proof.*

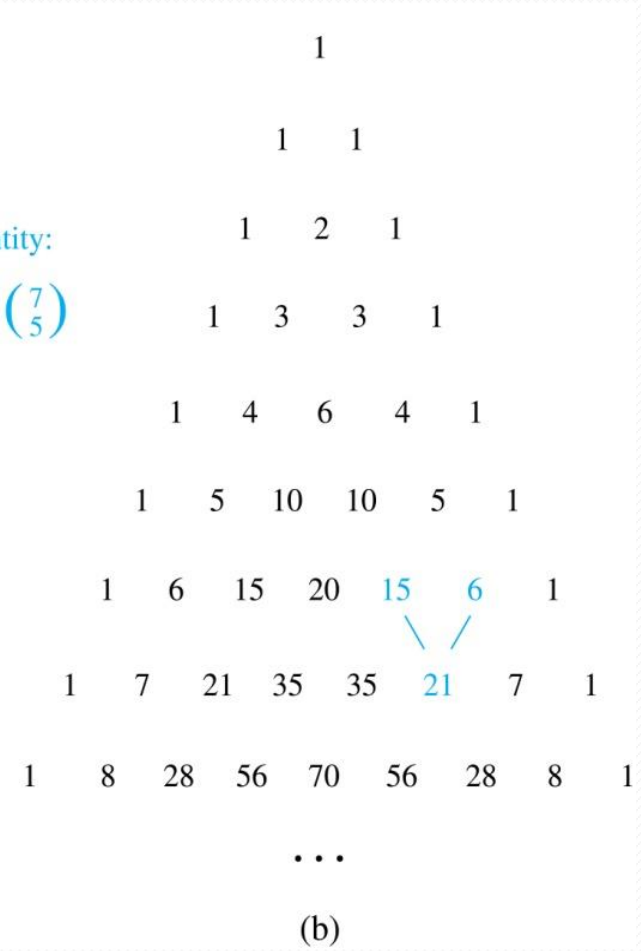
Pascal's Triangle

The nth row in the triangle consists of the binomial coefficients $\binom{n}{k}$, $k = 0,1,\ldots,n$.



By Pascal's identity:

$$\binom{6}{4} + \binom{6}{5} = \binom{7}{5}$$



By Pascal's identity, adding two adjacent bionomial coefficients results is the binomial coefficient in the next row between these two coefficients.

Graphs

Chapter 10

Chapter Summary

- Graphs and Graph Models
- Graph Terminology and Special Types of Graphs
- Representing Graphs and Graph Isomorphism
- Connectivity
- Euler and Hamiltonian Graphs
- Shortest-Path Problems
- Planar Graphs
- Graph Coloring

Graphs and Graph Models

Section 10.1

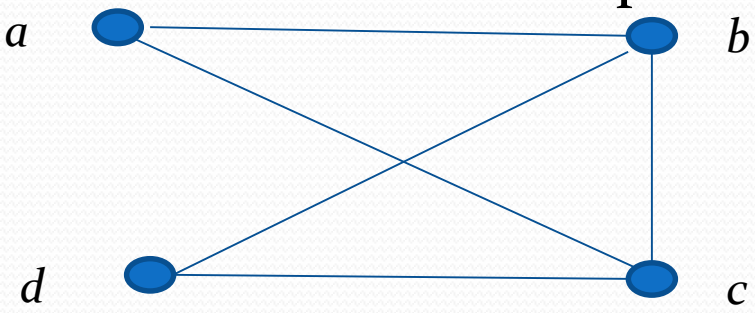
Section Summary

- Introduction to Graphs
- Graph Taxonomy
- Graph Models

Graphs

Definition: A *graph* $G = (V, E)$ consists of a nonempty set V of *vertices* (or *nodes*) and a set E of *edges*. Each edge has either one or two vertices associated with it, called its *endpoints*. An edge is said to *connect* its endpoints.

Example:



This is a graph with four vertices and five edges.

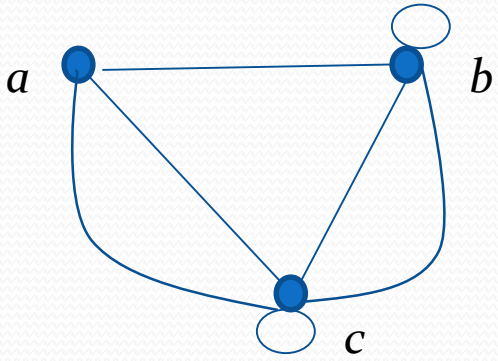
Remarks:

- The graphs we study here are unrelated to graphs of functions studied in Chapter 2.
- We have a lot of freedom when we draw a picture of a graph. All that matters is the connections made by the edges, not the particular geometry depicted. For example, the lengths of edges, whether edges cross, how vertices are depicted, and so on, do not matter
- A graph with an infinite vertex set is called an *infinite graph*. A graph with a finite vertex set is called a *finite graph*. We (following the text) restrict our attention to finite graphs.

Some Terminology

- In a *simple graph* each edge connects two different vertices and no two edges connect the same pair of vertices.
- *Multigraphs* may have multiple edges connecting the same two vertices. When m different edges connect the vertices u and v, we say that $\{u,v\}$ is an edge of *multiplicity* m.
- An edge that connects a vertex to itself is called a *loop*.
- A *pseudograph* may include loops, as well as multiple edges connecting the same pair of vertices.

Example:
This pseudograph has both multiple edges and a loop.



Remark: There is no standard terminology for graph theory. So, it is crucial that you understand the terminology being used whenever you read material about graphs.

Directed Graphs

Definition: An *directed graph* (or *digraph*) $G = (V, E)$ consists of a nonempty set V of *vertices* (or *nodes*) and a set E of *directed edges* (or *arcs*). Each 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.

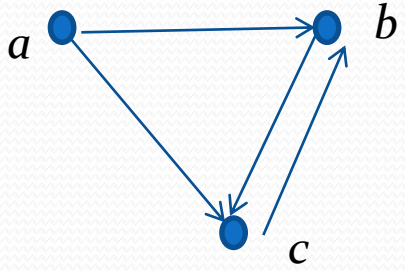
Remark:

- Graphs where the end points of an edge are not ordered are said to be *undirected graphs*.

Some Terminology (*continued*)

- A *simple directed graph* has no loops and no multiple edges.

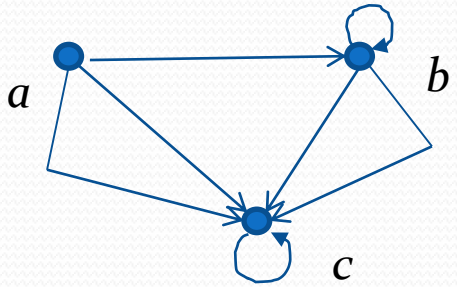
Example:
This is a directed graph with three vertices and four edges.



- A *directed multigraph* may have multiple directed edges. When there are m directed edges from the vertex u to the vertex v, we say that (u,v) is an edge of *multiplicity* m.

Example:

In this directed multigraph the multiplicity of (a,b) is 1 and the multiplicity of (b,c) is 2.



Graph Terminology: Summary

- To understand the structure of a graph and to build a graph model, we ask these questions:
  - Are the edges of the graph undirected or directed (or both)?
  - If the edges are undirected, are multiple edges present that connect the same pair of vertices? If the edges are directed, are multiple directed edges present?
  - Are loops present?

TABLE 1 Graph Terminology.

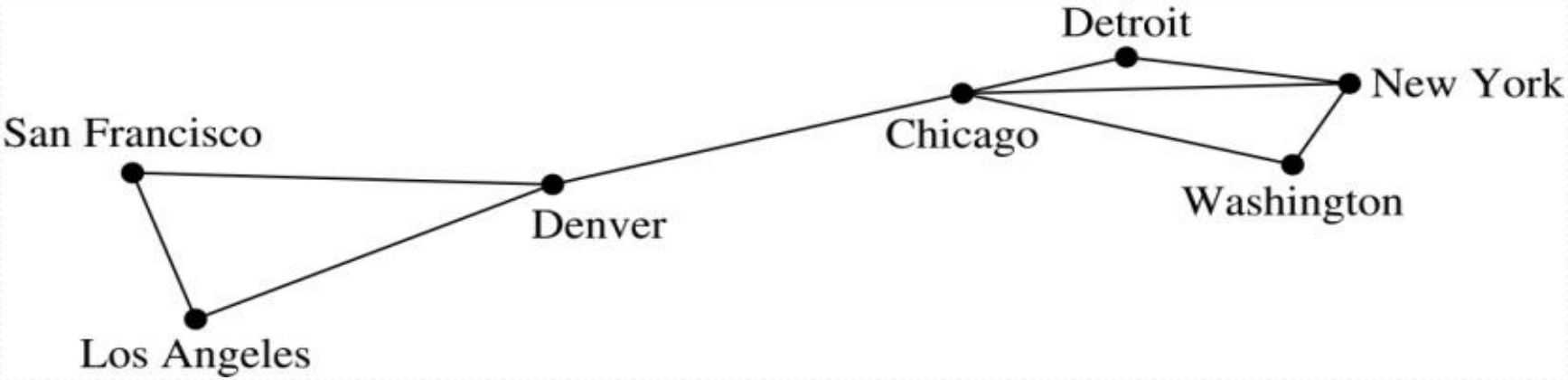
Type	*Edges*	*Multiple Edges Allowed?*	*Loops Allowed?*
Simple graph	Undirected	No	No
Multigraph	Undirected	Yes	No
Pseudograph	Undirected	Yes	Yes
Simple directed graph	Directed	No	No
Directed multigraph	Directed	Yes	Yes
Mixed graph	Directed and undirected	Yes	Yes

Other Applications of Graphs

- We will illustrate how graph theory can be used in models of:
  - Social networks
  - Communications networks
  - Information networks
  - Software design
  - Transportation networks
  - Biological networks
- It's a challenge to find a subject to which graph theory has not yet been applied. Can you find an area without applications of graph theory?

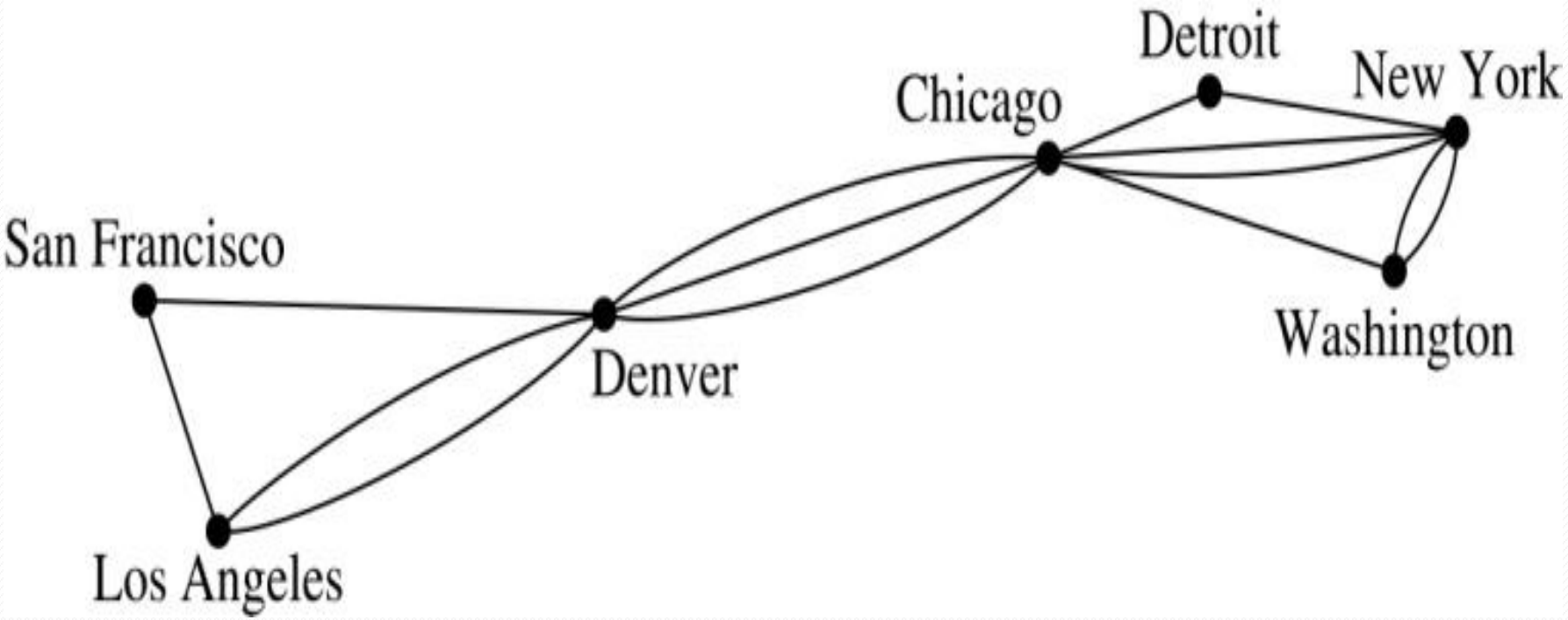
Graph Models: Computer Networks

- When we build a graph model, we use the appropriate type of graph to capture the important features of the application.
- We illustrate this process using graph models of different types of computer networks. In all these graph models, the vertices represent data centers and the edges represent communication links.
- To model a computer network where we are only concerned whether two data centers are connected by a communications link, we use a simple graph. This is the appropriate type of graph when we only care whether two data centers are directly linked (and not how many links there may be) and all communications links work in both directions.



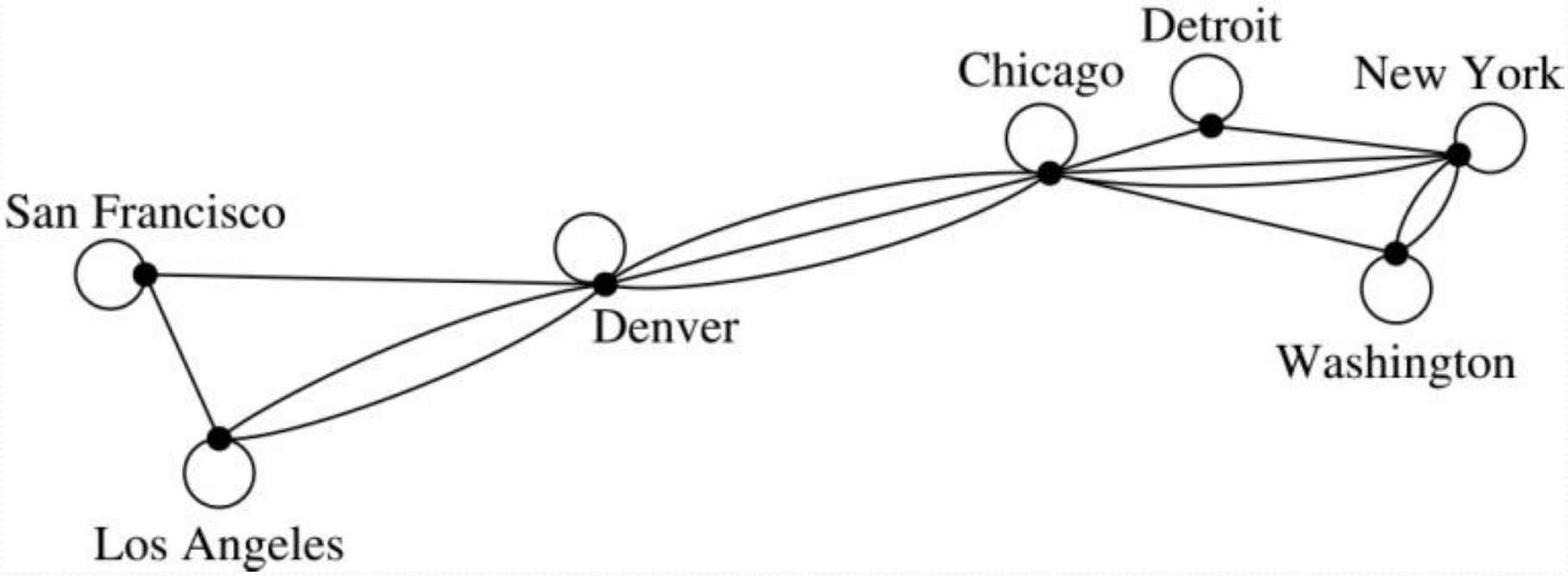
Graph Models: Computer Networks (*continued*)

- To model a computer network where we care about the number of links between data centers, we use a multigraph.



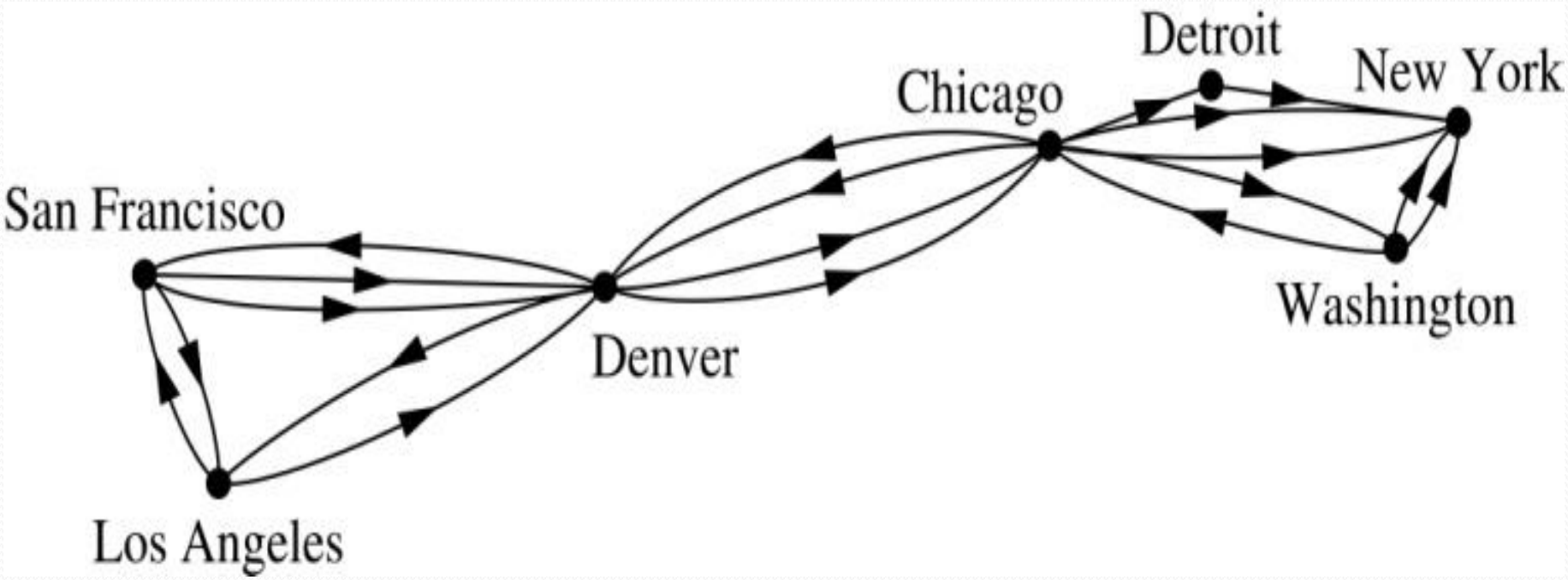
Graph Models: Computer Networks

- To model a computer network with diagnostic links at data centers, we use a pseudograph, as loops are needed.



Graph Models: Computer Networks

- To model a network with multiple one-way links, we use a directed multigraph. Note that we could use a directed graph without multiple edges if we only care whether there is at least one link from a data center to another data center.

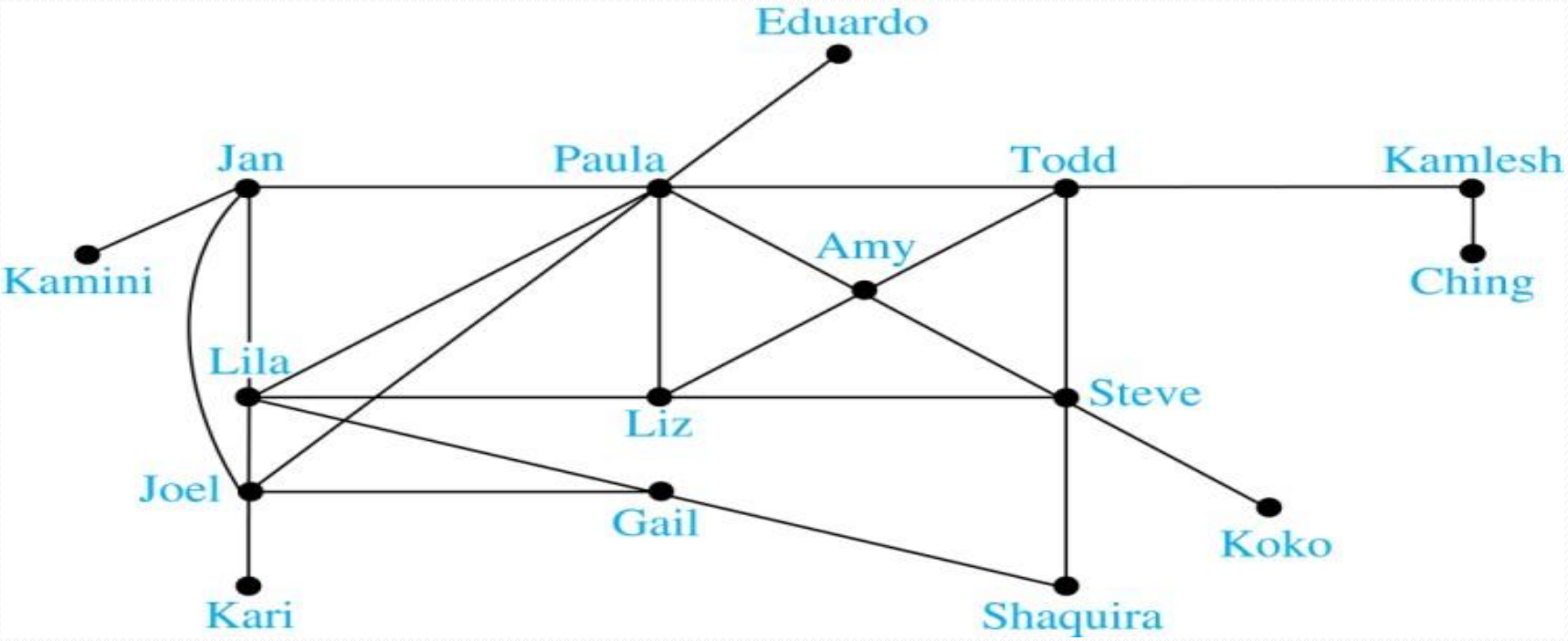


Graph Models: Social Networks

- Graphs can be used to model social structures based on different kinds of relationships between people or groups.
- In a *social network*, vertices represent individuals or organizations and edges represent relationships between them.
- Useful graph models of social networks include:
  - *friendship graphs* - undirected graphs where two people are connected if they are friends (in the real world, on Facebook, or in a particular virtual world, and so on.)
  - *collaboration graphs* - undirected graphs where two people are connected if they collaborate in a specific way
  - *influence graphs* - directed graphs where there is an edge from one person to another if the first person can influence the second person

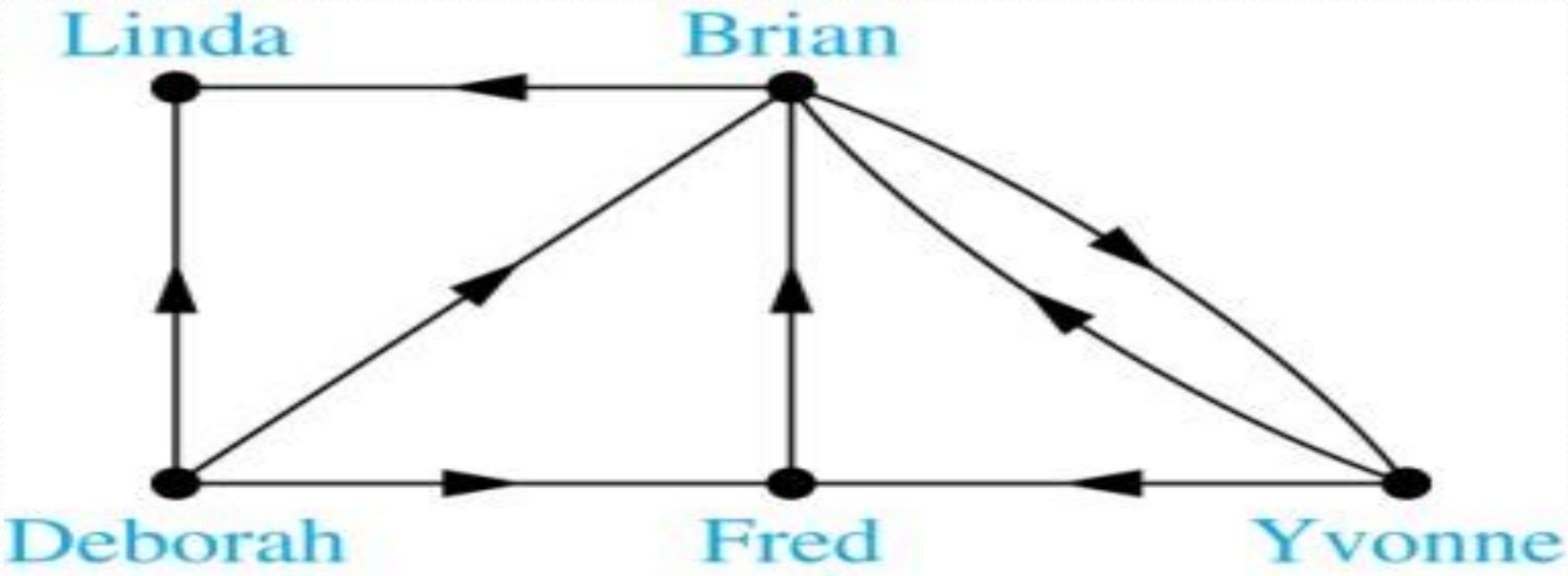
Graph Models: Social Networks

Example: A friendship graph where two people are connected if they are Facebook friends.



Graph Models: Social Networks

Example: An influence graph



Applications to Information Networks

- Graphs can be used to model different types of networks that link different types of information.
- In a *web graph*, web pages are represented by vertices and links are represented by directed edges.
  - A web graph models the web at a particular time.
  - We will explain how the web graph is used by search engines in Section 11.4.
- In a *citation network*:
  - Research papers in a particular discipline are represented by vertices.
  - When a paper cites a second paper as a reference, there is an edge from the vertex representing this paper to the vertex representing the second paper.

Transportation Graphs

- Graph models are extensively used in the study of transportation networks.
- Airline networks can be modeled using directed multigraphs where
  - airports are represented by vertices
  - each flight is represented by a directed edge from the vertex representing the departure airport to the vertex representing the destination airport
- Road networks can be modeled using graphs where
  - vertices represent intersections and edges represent roads.
  - undirected edges represent two-way roads and directed edges represent one-way roads.

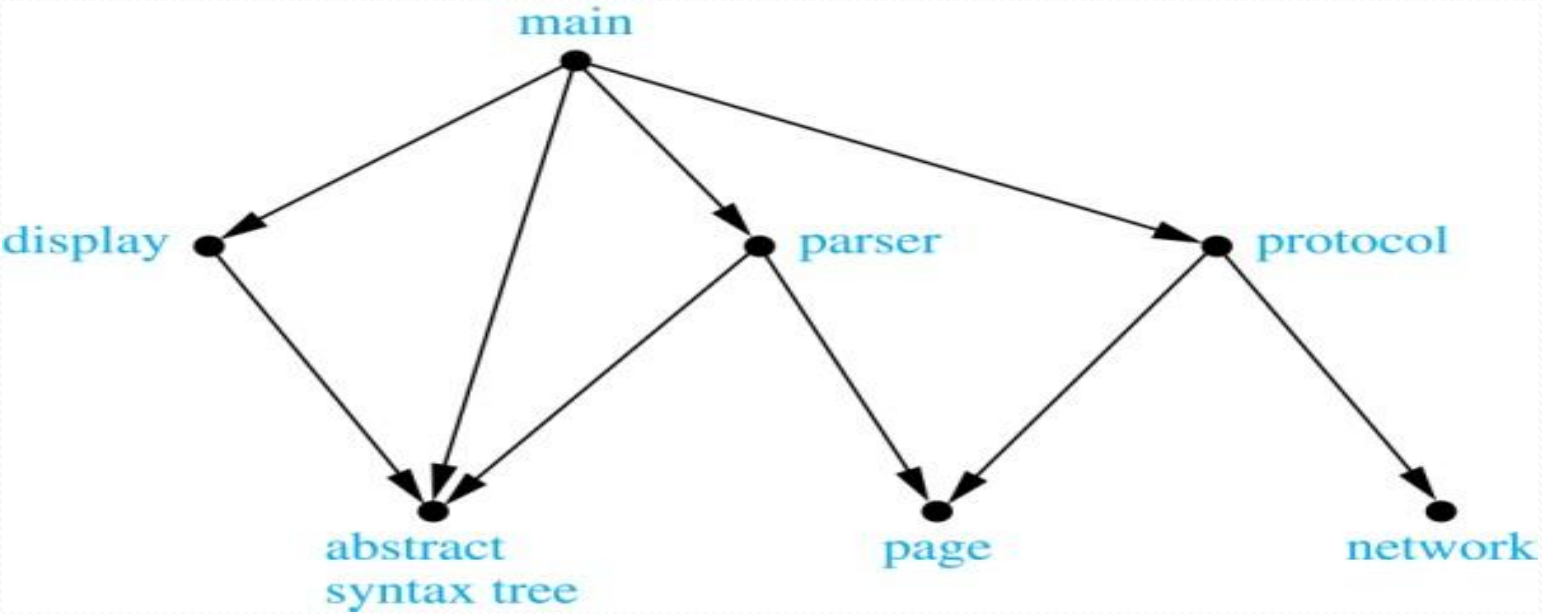
Software Design Applications

- Graph models are extensively used in software design. We will introduce two such models here; one representing the dependency between the modules of a software application and the other representing restrictions in the execution of statements in computer programs.
- When a top-down approach is used to design software, the system is divided into modules, each performing a specific task.
- We use a *module dependency graph* to represent the dependency between these modules. These dependencies need to be understood before coding can be done.

Software Design Applications

- In a module dependency graph vertices represent software modules and there is an edge from one module to another if the second module depends on the first.

Example: The dependencies between the seven modules in the design of a web browser are represented by this module dependency graph.



Software Design Applications

- We can use a directed graph called a *precedence graph* to represent which statements must have already been executed before we execute each statement.
  - Vertices represent statements in a computer program
  - There is a directed edge from a vertex to a second vertex if the second vertex cannot be executed before the first

Example: This precedence graph shows which statements must already have been executed before we can execute each of the six statements in the program.

S_1 $a := 0$

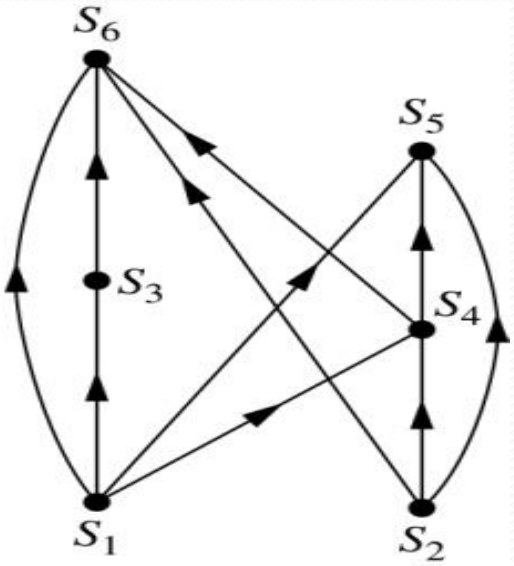
S_2 $b := 1$

S_3 $c := a + 1$

S_4 $d := b + a$

S_5 $e := d + 1$

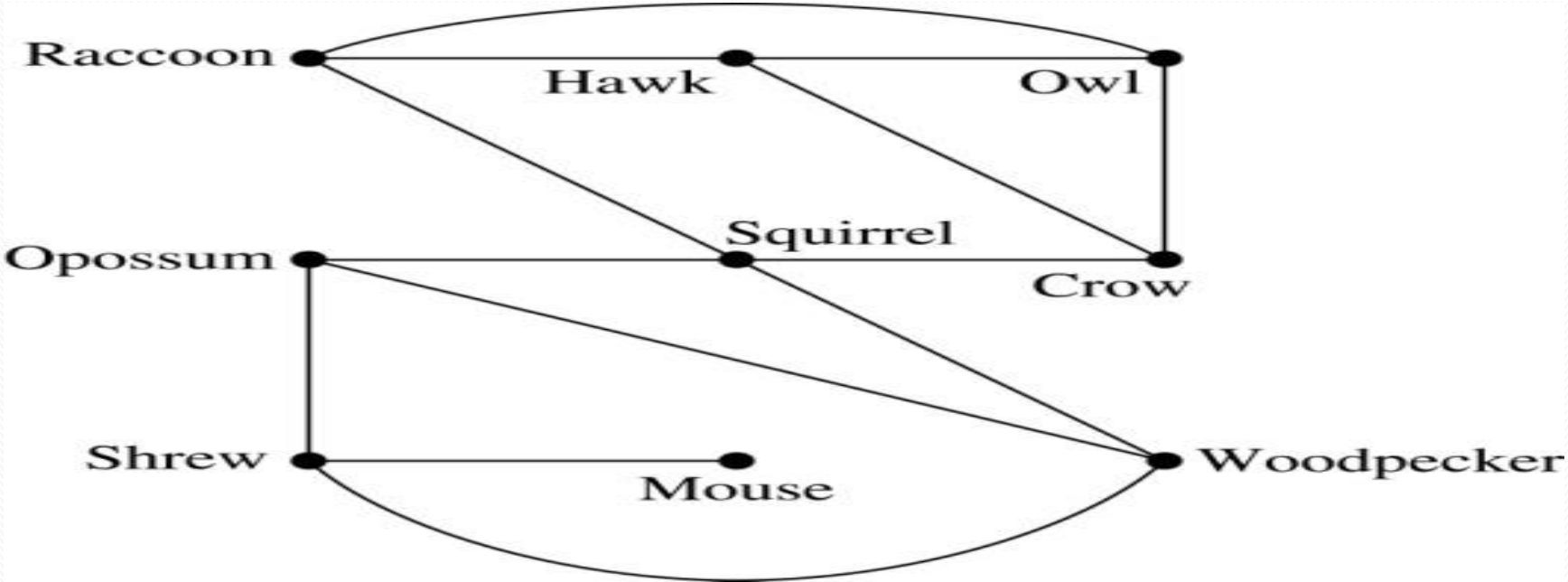
S_6 $e := c + d$



Biological Applications

- Graph models are used extensively in many areas of the biological science. We will describe two such models, one to ecology and the other to molecular biology.
- *Niche overlap graphs* model competition between species in an ecosystem
  - Vertices represent species and an edge connects two vertices when they represent species who compete for food resources.

Example: This is the niche overlap graph for a forest ecosystem with nine species.

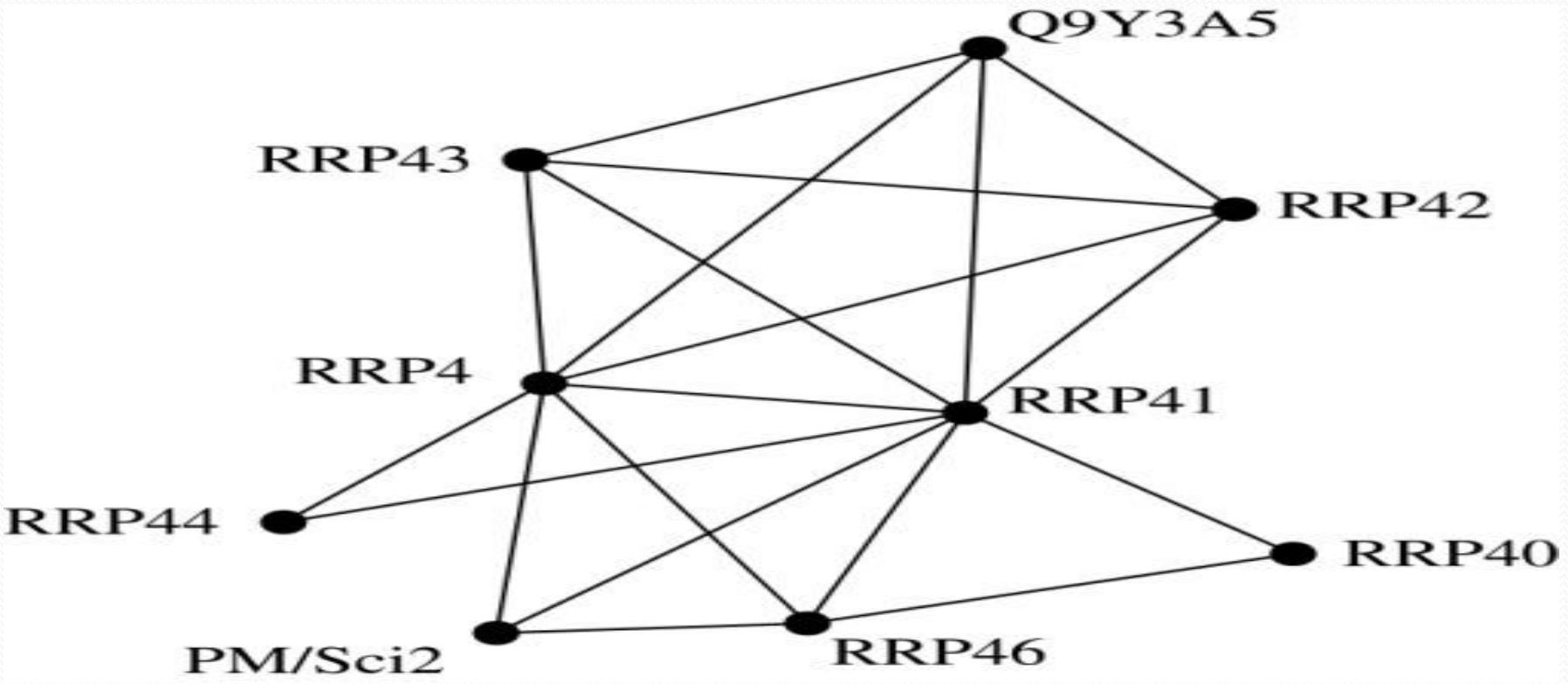


Biological Applications

- We can model the interaction of proteins in a cell using a *protein interaction network*.
- In a *protein interaction graph*, vertices represent proteins and vertices are connected by an edge if the proteins they represent interact.
- Protein interaction graphs can be huge and can contain more than 100,000 vertices, each representing a different protein, and more than 1,000,000 edges, each representing an interaction between proteins
- Protein interaction graphs are often split into smaller graphs, called *modules*, which represent the interactions between proteins involved in a particular function.

Biological Applications

Example: This is a module of the protein interaction graph of proteins that degrade RNA in a human cell.



Graph Terminology and Special Types of Graphs

Section 10.2

Section Summary

- Basic Terminology
- Some Special Types of Graphs
- Bipartite Graphs
- Bipartite Graphs and Matchings
- Some Applications of Special Types of Graphs
- New Graphs from Old

Basic Terminology

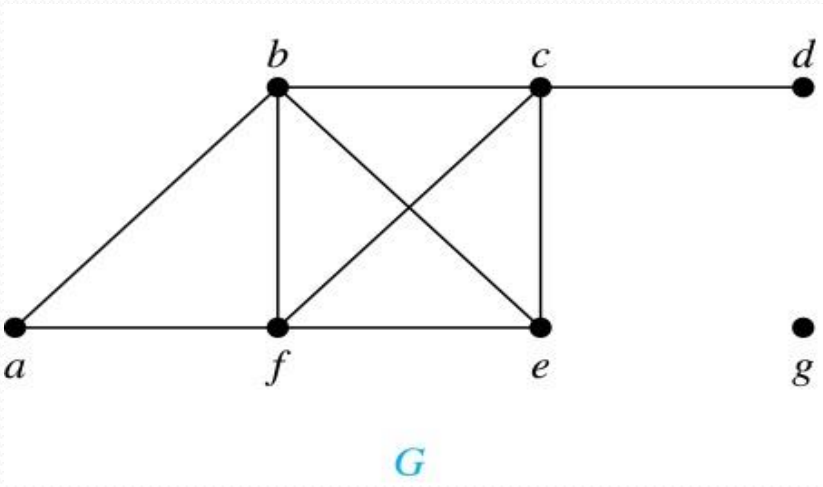
Definition 1. Two vertices *u*, *v* in an undirected graph *G* are called *adjacent* (or *neighbors*) in *G* if there is an edge *e* between *u* and *v*. Such an edge *e* is called *incident with* the vertices *u* and *v* and *e* is said to *connect u* and *v*.

Definition 2. 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, $N(A) = \bigcup_{v \in A} N(v)$.

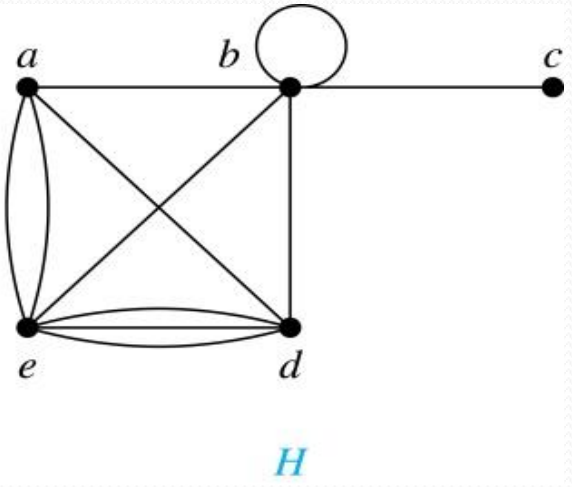
Definition 3. The *degree of a vertex in a undirected graph* is the number of edges incident with it, except that a loop at a vertex contributes two to the degree of that vertex. The degree of the vertex *v* is denoted by $\deg(v)$.

Degrees and Neighborhoods of Vertices

Example: What are the degrees and neighborhoods of the vertices in the graphs G and H?



G



H

Solution:

G: $\deg(a) = 2, \deg(b) = \deg(c) = \deg(f) = 4, \deg(d) = 1, \deg(e) = 3, \deg(g) = 0.$

$N(a) = \{b, f\}, N(b) = \{a, c, e, f\}, N(c) = \{b, d, e, f\}, N(d) = \{c\},$

$N(e) = \{b, c, f\}, N(f) = \{a, b, c, e\}, N(g) = \varnothing$.

H: $\deg(a) = 4, \deg(b) = \deg(e) = 6, \deg(c) = 1, \deg(d) = 5.$

$N(a) = \{b, d, e\}, N(b) = \{a, b, c, d, e\}, N(c) = \{b\}, N(d) = \{a, b, e\},$

$N(e) = \{a, b, d\}.$

Handshaking Theorem

Theorem: If G is any graph, then the sum of the degrees of all the vertices of G equals twice the number of edges of G.

Specifically, if the vertices of G are **v1**, **v2**, …, **vn**, where n is a positive integer, then

the total degree of G = deg(v1) + deg(v2) + … + deg(vn)
= 2 . (the number of edges of G)

PROOF:

- Each edge "e" of G connects its end points vi and vj. This edge, therefore contributes 1 to the degree of vi and 1 to the degree of vj.
- If "e" is a loop, then it is counted twice in computing the degree of the vertex on which it is incident.
- Accordingly, each edge of G contributes 2 to the total degree of G.

Thus, the total degree of G = 2. (the number of edges of G)

COROLLARY: The total degree of G is an even number

Degree of Vertices

Theorem: An undirected graph has an even number of vertices of odd degree.

***Proof*:** Let V_1 be the vertices of even degree and V_2 be the vertices of odd degree in an undirected graph $G = (V, E)$ with m edges. Then

even → $$2m = \sum_{v \in V} \deg(v) = \sum_{v \in V_1} \deg(v) + \sum_{v \in V_2} \deg(v).$$

must be even since deg(v) is even for each $v \in V_1$

This sum must be even because $2m$ is even and the sum of the degrees of the vertices of even degrees is also even. Because this is the sum of the degrees of all vertices of odd degree in the graph, there must be an even number of such vertices.

Handshaking Theorem

We now give two examples illustrating the usefulness of the handshaking theorem.

Example: How many edges are there in a graph with 10 vertices of degree six?

Solution: Because the sum of the degrees of the vertices is $6 \cdot 10 = 60$, the handshaking theorem tells us that $2m = 60$. So the number of edges $m = 30$.

Example: If a graph has 5 vertices, can each vertex have degree 3?

Solution: This is not possible by the handshaking theorem, because the sum of the degrees of the vertices $3 \cdot 5 = 15$ is odd.

Directed Graphs

Recall the definition of a directed graph.

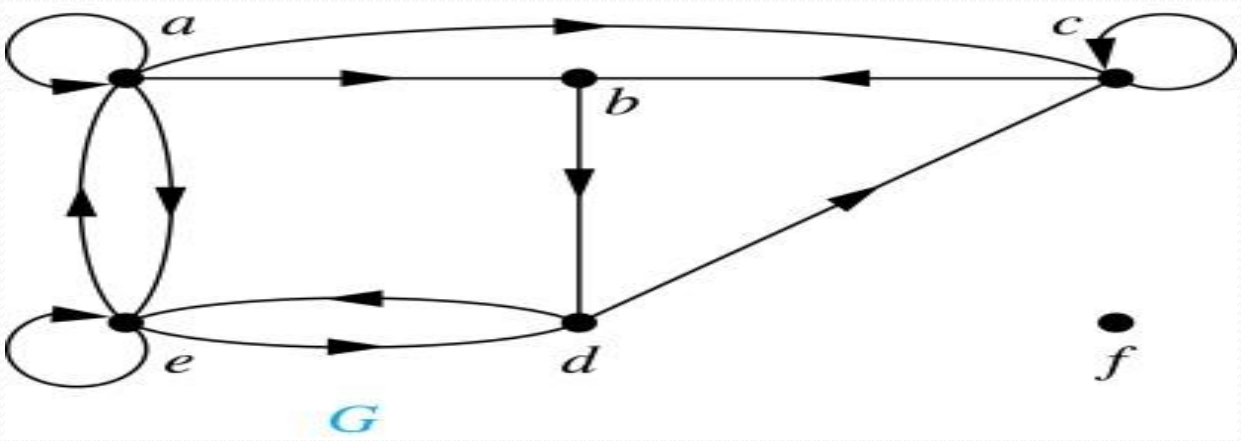
Definition: An *directed graph* $G = (V, E)$ consists of V, a nonempty set of *vertices* (or *nodes*), and E, a set of *directed edges* or *arcs*. Each edge is an ordered pair of vertices. The directed edge (u,v) is said to start at u and end at v.

Definition: Let (u,v) be an edge in G. Then u is the *initial vertex* of this edge and is *adjacent to* v and v is the *terminal* (or *end*) *vertex* of this edge and is *adjacent from* u. The initial and terminal vertices of a loop are the same.

Directed Graphs (*continued*)

Definition: The *in-degree of a vertex v*, denoted $deg^-(v)$, is the number of edges which terminate at *v*. The *out-degree of v*, denoted $deg^+(v)$, is the number of edges with *v* as their initial vertex. Note that a loop at a vertex contributes 1 to both the in-degree and the out-degree of the vertex.

Example: In the graph *G* we have



$\deg^-(a) = 2, \deg^-(b) = 2, \deg^-(c) = 3, \deg^-(d) = 2, \deg^-(e) = 3, \deg^-(f) = 0.$

$\deg^+(a) = 4, \deg^+(b) = 1, \deg^+(c) = 2, \deg^+(d) = 2, \deg^+(e) = 3, \deg^+(f) = 0.$

Directed Graphs (*continued*)

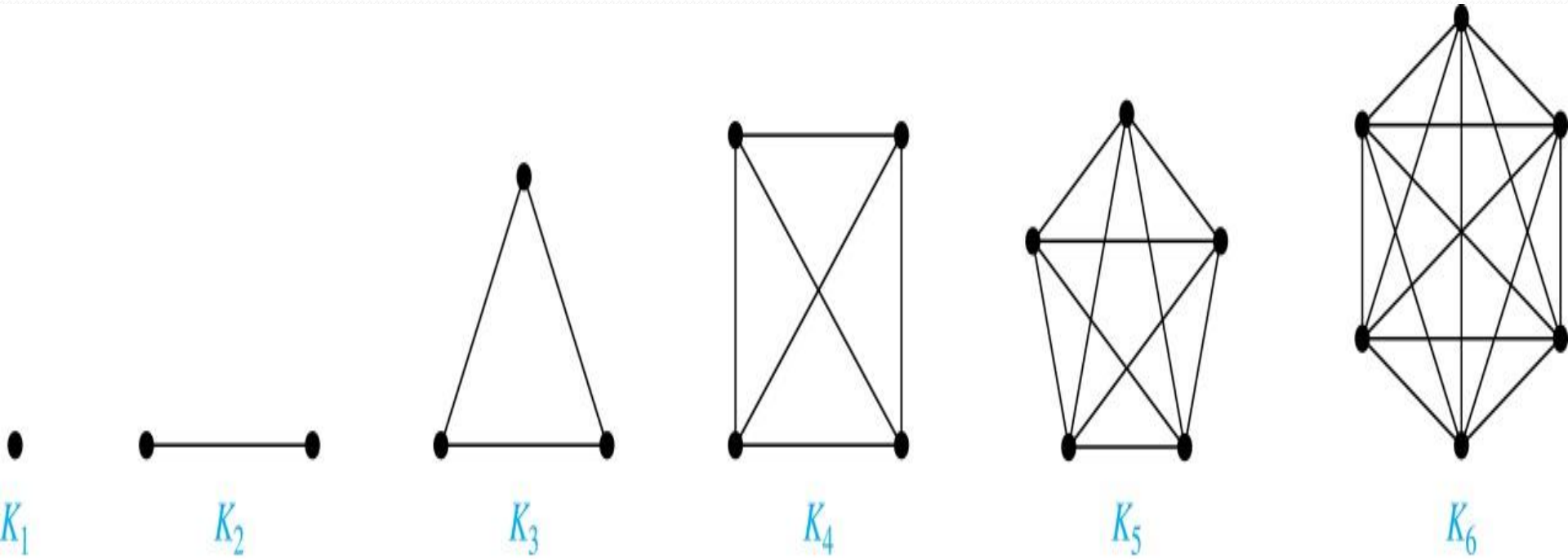
Theorem 3: Let $G = (V, E)$ be a graph with directed edges. Then:

$$|E| = \sum_{v \in V} deg^-(v) = \sum_{v \in V} deg^+(v).$$

Proof: The first sum counts the number of outgoing edges over all vertices and the second sum counts the number of incoming edges over all vertices. It follows that both sums equal the number of edges in the graph.

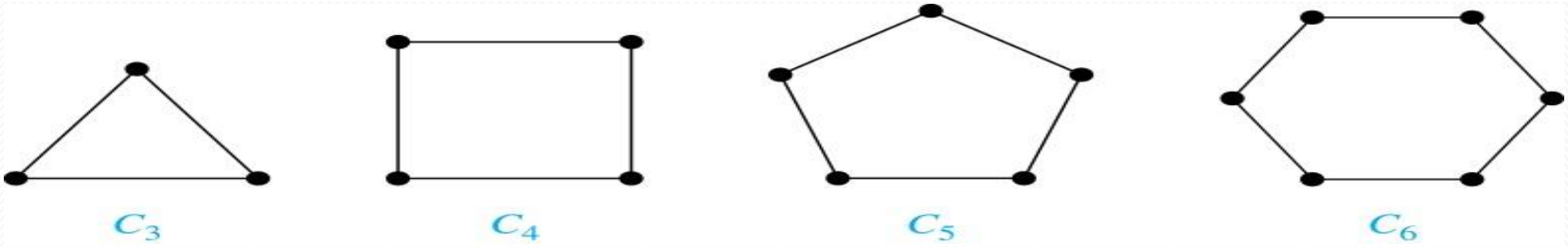
Special Types of Simple Graphs: Complete Graphs

A *complete graph on n vertices*, denoted by K_n, is the simple graph that contains exactly one edge between each pair of distinct vertices.

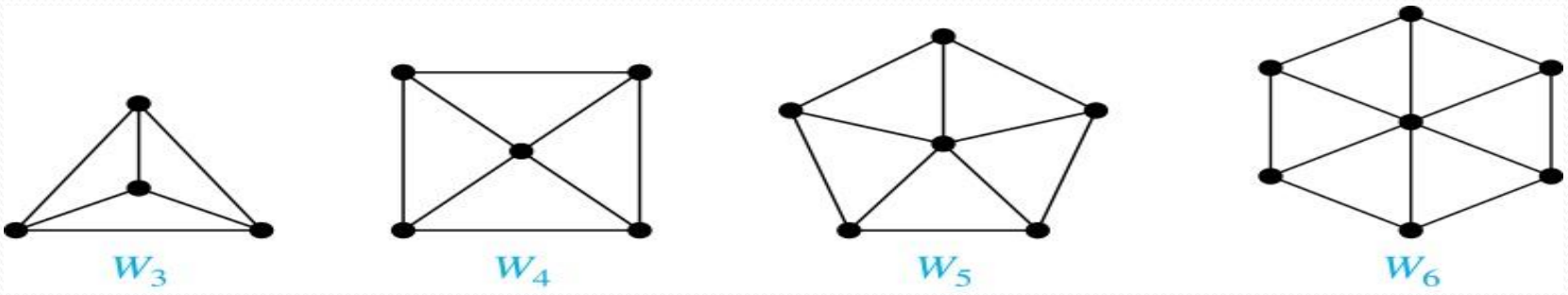


Special Types of Simple Graphs: Cycles and Wheels

A *cycle* C_n for $n \geq 3$ consists of n vertices $v_1, v_2, \cdots, v_n$, and edges $\{v_1, v_2\}, \{v_2, v_3\}, \cdots, \{v_{n-1}, v_n\}, \{v_n, v_1\}$.

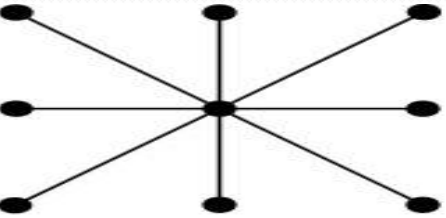


A *wheel* W_n is obtained by adding an additional vertex to a cycle C_n for $n \geq 3$ and connecting this new vertex to each of the n vertices in C_n by new edges.

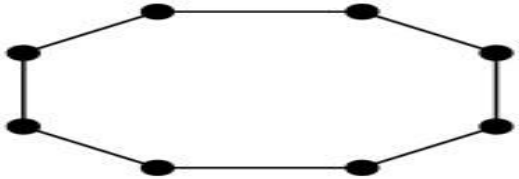


Special Types of Graphs and Computer Network Architecture

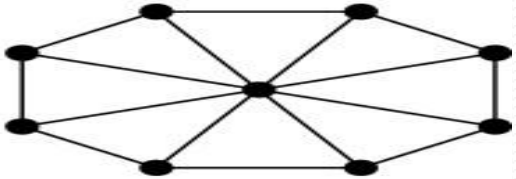
Various special graphs play an important role in the design of computer networks.



(a)



(b)

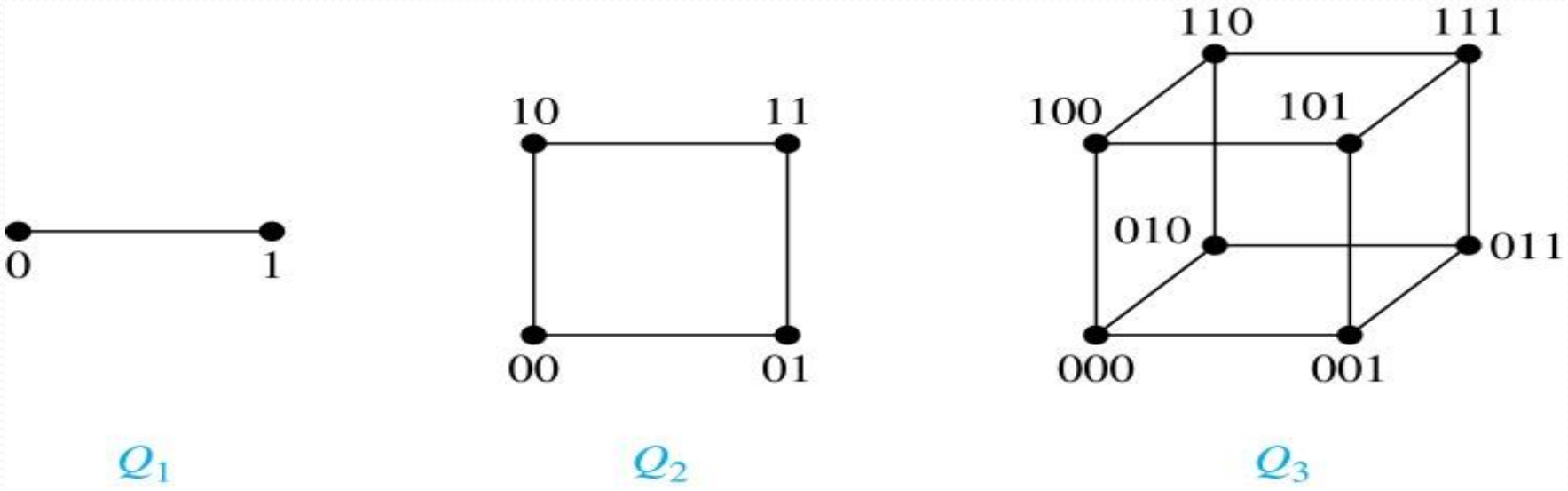


(c)

- Some local area networks use a *star topology*, which is a complete bipartite graph $K_{1,n}$, as shown in (a). All devices are connected to a central control device.
- Other local networks are based on a *ring topology*, where each device is connected to exactly two others using C_n, as illustrated in (b). Messages may be sent around the ring.
- Others, as illustrated in (c), use a W_n – based topology, combining the features of a star topology and a ring topology.

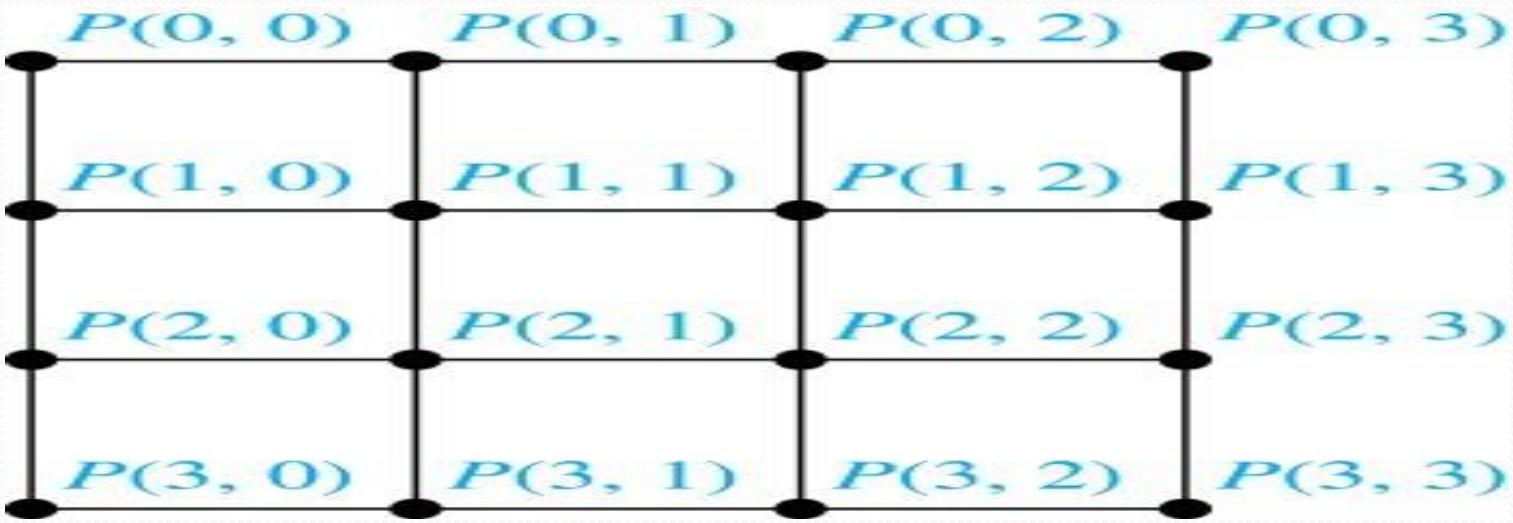
Special Types of Simple Graphs: *n*-Cubes

An *n-dimensional hypercube*, or *n-cube*, $\boldsymbol{Q_n}$, is a graph with 2^n vertices representing all bit strings of length n, where there is an edge between two vertices that differ in exactly one bit position.



Special Types of Graphs and Computer Network Architecture

- Various special graphs also play a role in parallel processing where processors need to be interconnected as one processor may need the output generated by another.
- The n-dimensional hypercube, or n-cube, Qn, is a common way to connect processors in parallel, e.g., Intel Hypercube.
- Another common method is the mesh network, illustrated here for 16 processors.



Bipartite Graphs

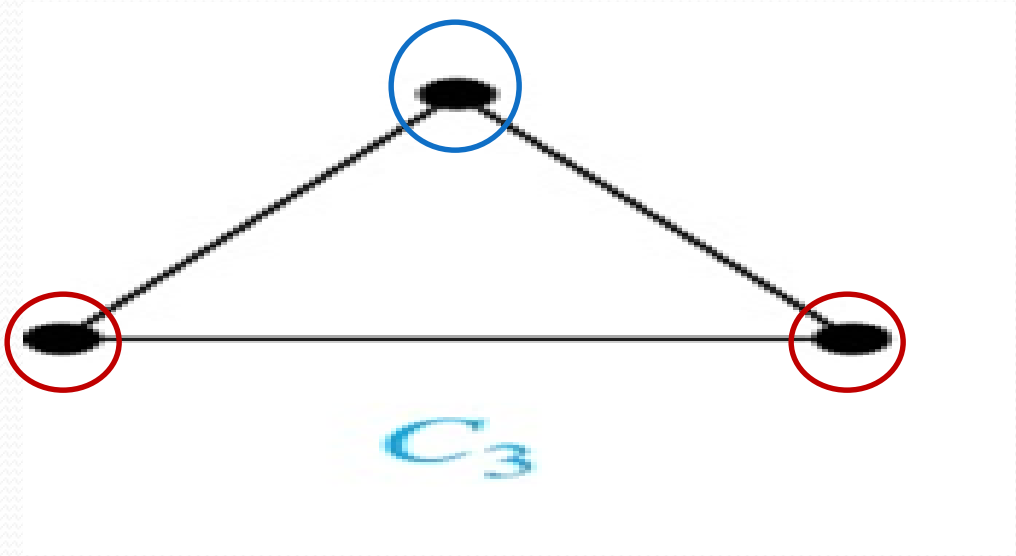
Definition: A simple graph G is bipartite if V can be partitioned into two disjoint subsets V_1 and V_2 such that every edge connects a vertex in V_1 and a vertex in V_2. In other words, there are no edges which connect two vertices in V_1 or in V_2.

- It is not hard to show that an equivalent definition of a bipartite graph is a graph where it is possible to color the vertices red or blue so that no two adjacent vertices are the same color.

Bipartite Graphs (*continued*)

Example: Show that C_3 is not bipartite.

Solution: If we divide the vertex set of C_3 into two nonempty sets, one of the two must contain two vertices. But in C_3 every vertex is connected to every other vertex. Therefore, the two vertices in the same partition are connected. Hence, C_3 is not bipartite.

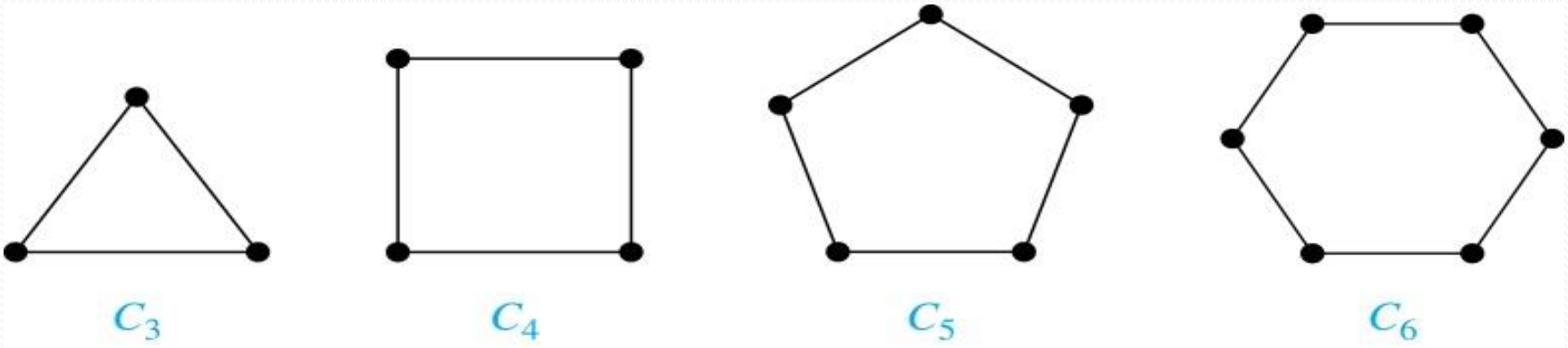


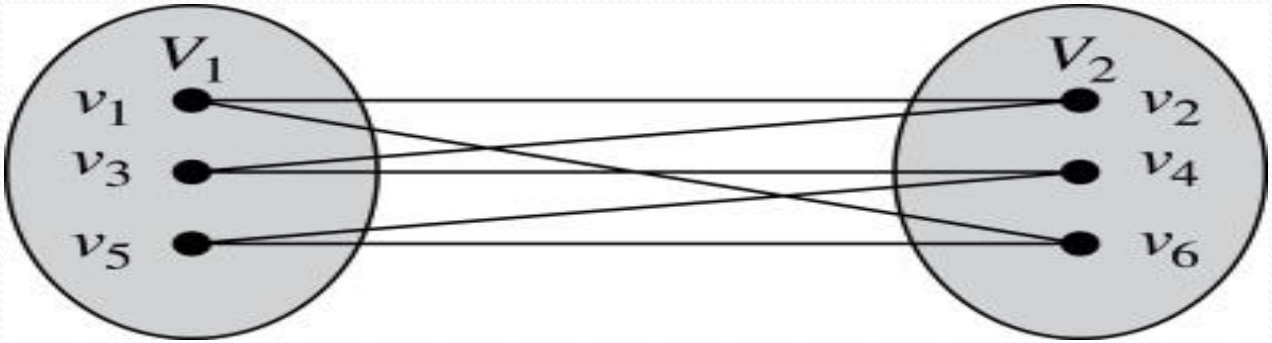
Bipartite Graphs (*continued*)

Example: Show that C_6 is bipartite.

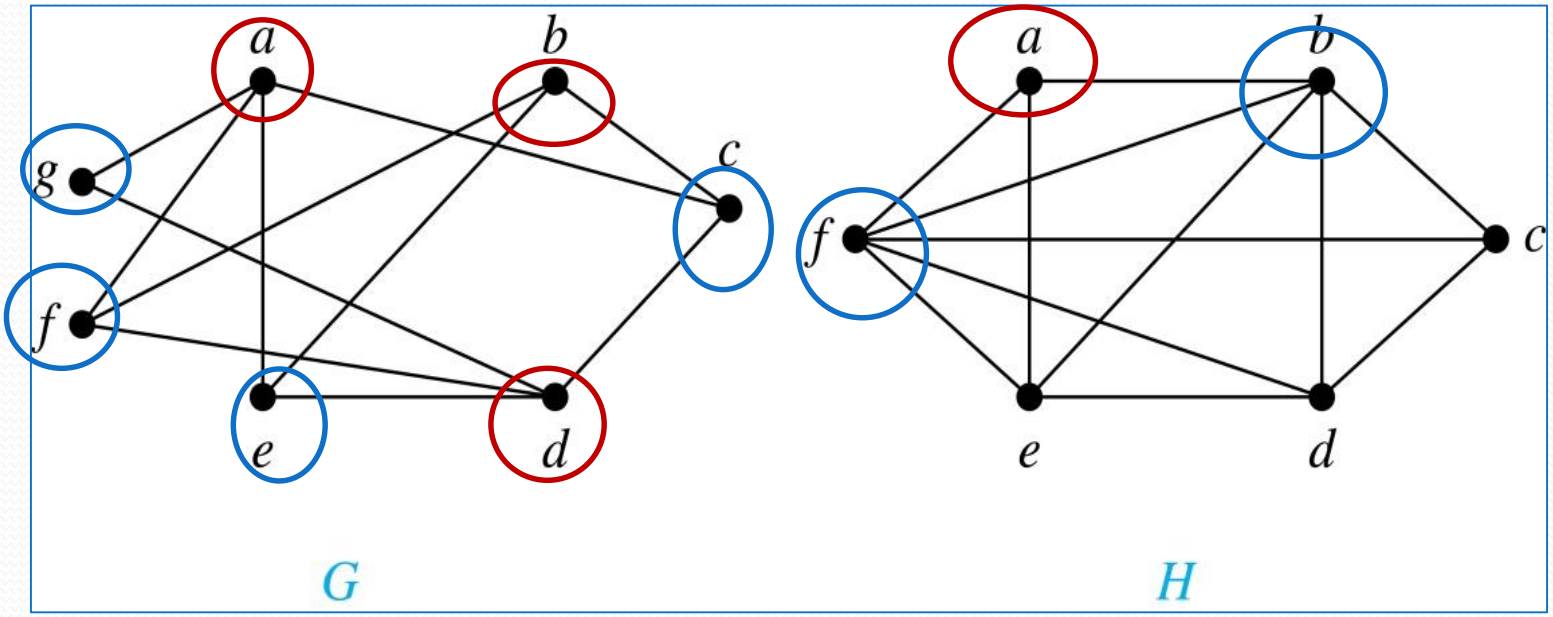
Solution: We can partition the vertex set into

$V_1 = \{v_1, v_3, v_5\}$ and $V_2 = \{v_2, v_4, v_6\}$ so that every edge of C_6 connects a vertex in V_1 and V_2 .





Bipartite Graphs

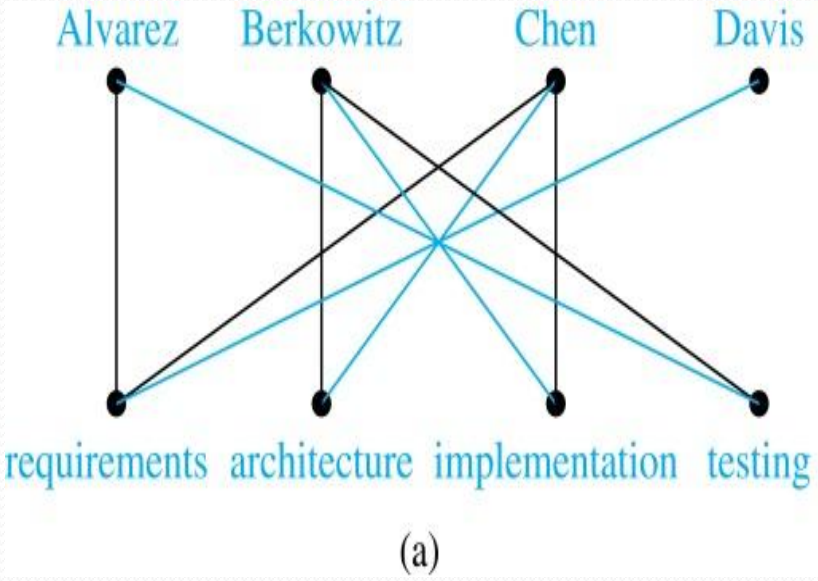


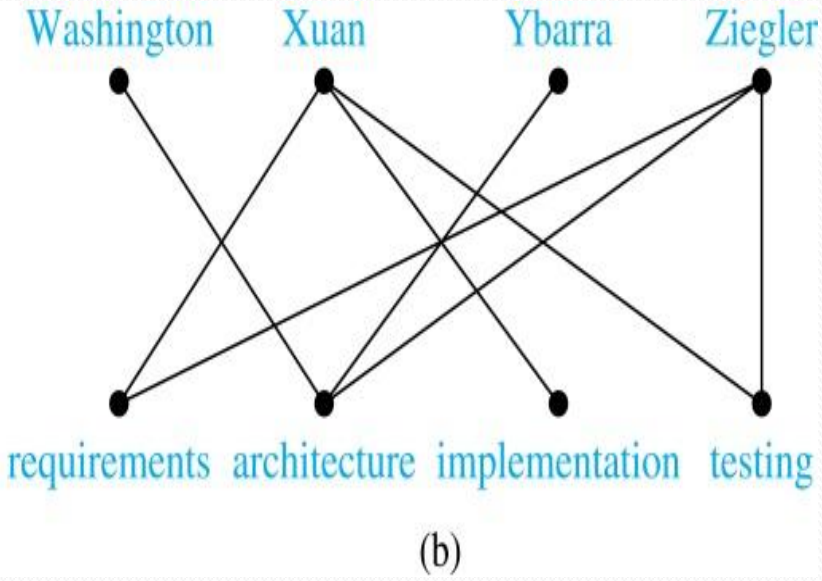
G is bipartite

H is not bipartite since if we color *a* red, then the adjacent vertices *f* and *b* must both be blue.

Bipartite Graphs and Matchings

- Bipartite graphs are used to model applications that involve matching the elements of one set to elements in another, for example:
- *Job assignments* - vertices represent the jobs and the employees, edges link employees with those jobs they have been trained to do. A common goal is to match jobs to employees so that the most jobs are done.

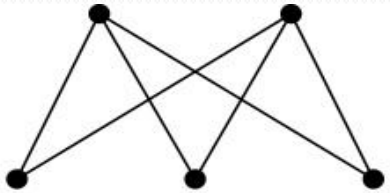




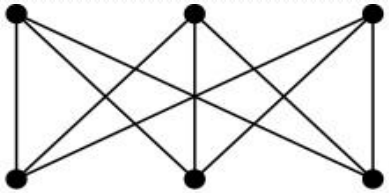
Complete Bipartite Graphs

Definition: A *complete bipartite graph* $K_{m,n}$ is a graph that has its vertex set partitioned into two subsets V_1 of size m and V_2 of size n such that there is an edge from every vertex in V_1 to every vertex in V_2.

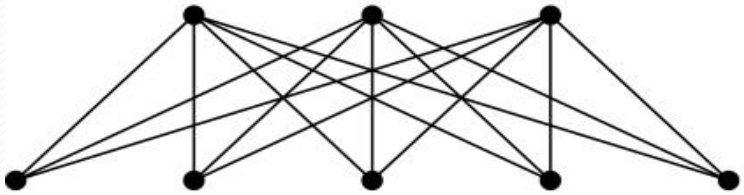
Example: We display four complete bipartite graphs here.



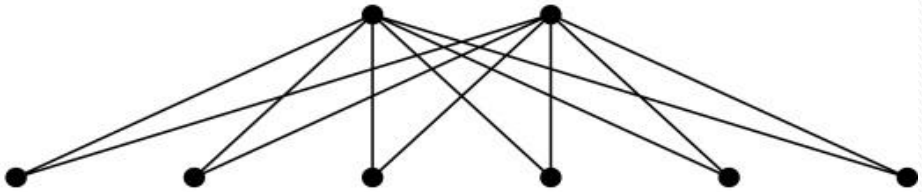
$K_{2,3}$



$K_{3,3}$



$K_{3,5}$

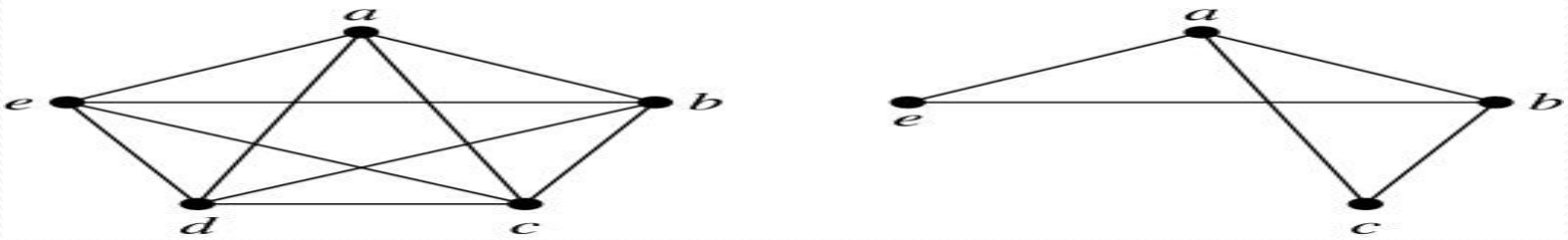


$K_{2,6}$

New Graphs from Old

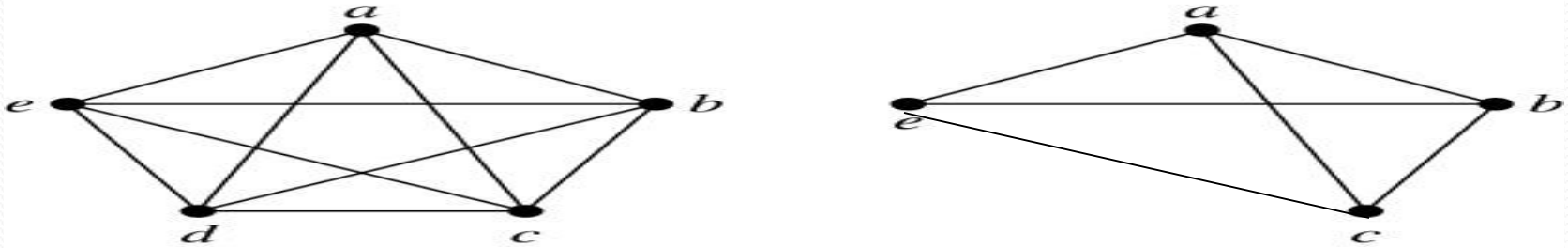
Definition: A *subgraph of a graph* $G = (V,E)$ is a graph (W,F), where $W \subset V$ and $F \subset E$. A subgraph H of G is a proper subgraph of G if $H \neq G$.

Example: Here we show K_5 and one of its subgraphs.



Definition: Let $G = (V, E)$ be a simple graph. The *subgraph induced* by a subset W of the vertex set V is the graph (W,F), where the edge set F contains an edge in E if and only if both endpoints are in W.

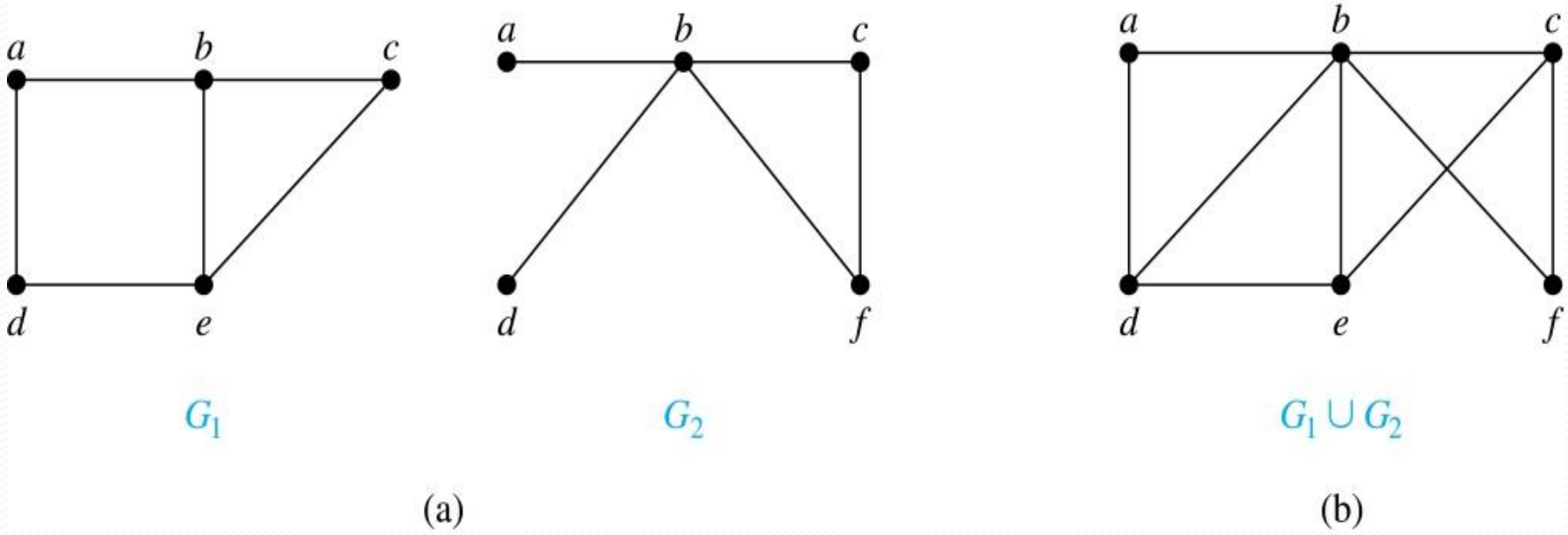
Example: Here we show K_5 and the subgraph induced by $W = \{a,b,c,e\}$.



New Graphs from Old (*continued*)

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:



Representations of Graphs

Section 10.3

Section Summary

- Adjacency Lists
- Adjacency Matrices
- Incidence Matrices

Representing Graphs: Adjacency Lists

Definition: An *adjacency list* can be used to represent a graph with no multiple edges by specifying the vertices that are adjacent to each vertex of the graph.

Example:

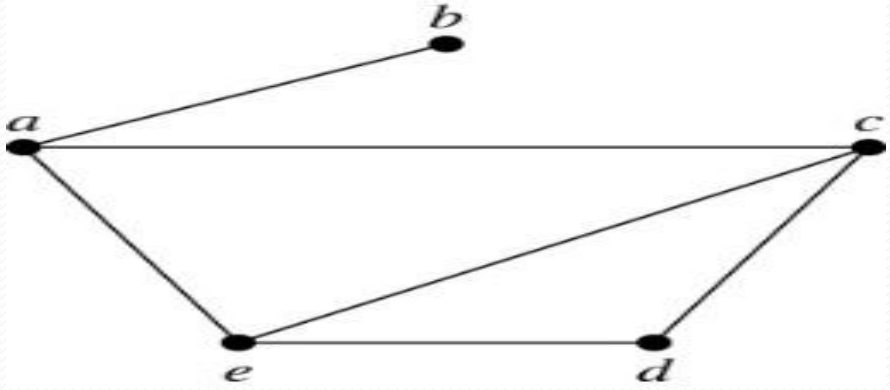


TABLE 1 An Adjacency List for a Simple Graph.

Vertex	*Adjacent Vertices*
a	b, c, e
b	a
c	a, d, e
d	c, e
e	a, c, d

Representing Graphs: Adjacency Lists

Example:

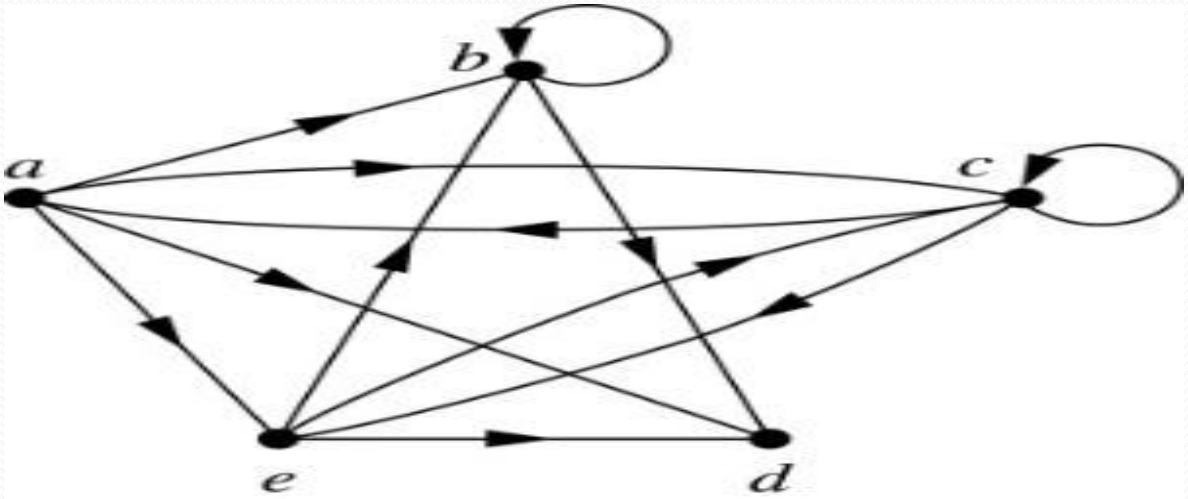


TABLE 2 An Adjacency List for a Directed Graph.

Initial Vertex	*Terminal Vertices*
a	b, c, d, e
b	b, d
c	a, c, e
d	
e	b, c, d

Representation of Graphs: Adjacency Matrices

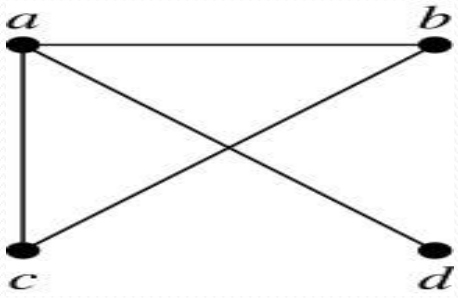
Definition: Suppose that $G = (V, E)$ is a simple graph where $|V| = n$. Arbitrarily list the vertices of G as $v_1, v_2, \ldots, v_n$. The *adjacency matrix* $\mathbf{A}_G$ of G, with respect to the listing of vertices, is the $n \times n$ zero-one matrix with 1 as its (i, j)th entry when v_i and v_j are adjacent, and 0 as its (i, j)th entry when they are not adjacent.

- In other words, if the graphs adjacency matrix is $\mathbf{A}_G = [a_{ij}]$, then

$$a_{ij} = \begin{cases} 1 & \text{if } \{v_i, v_j\} \text{ is an edge of } G, \\ 0 & \text{otherwise.} \end{cases}$$

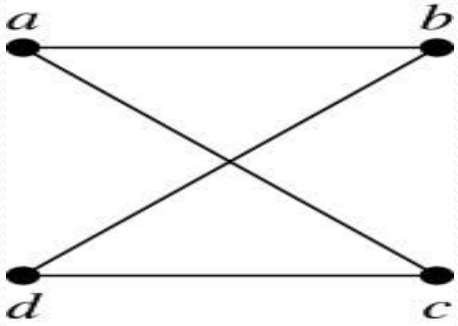
Adjacency Matrices (*continued*)

Example:



$$\begin{bmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}$$

The ordering of vertices is a, b, c, d.



$$\begin{bmatrix} 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \end{bmatrix}$$

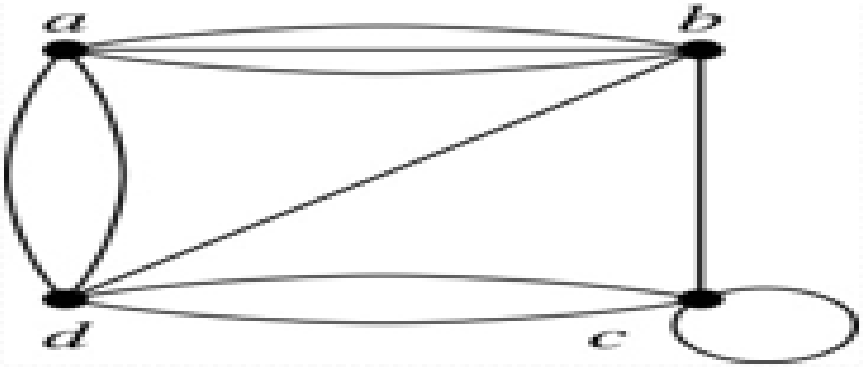
The ordering of vertices is a, b, c, d.

When a graph is sparse, that is, it has few edges relatively to the total number of possible edges, it is much more efficient to represent the graph using an adjacency list than an adjacency matrix. But for a dense graph, which includes a high percentage of possible edges, an adjacency matrix is preferable.

Note: The adjacency matrix of a simple graph is symmetric, i.e., $a_{ij} = a_{ji}$
Also, since there are no loops, each diagonal entry a_{ij} for $i = 1, 2, 3, \ldots, n$, is 0.

Adjacency Matrices (*continued*)

- Adjacency matrices can also be used to represent graphs with loops and multiple edges.
- A loop at the vertex vi is represented by a 1 at the (i, j)th position of the matrix.
- When multiple edges connect the same pair of vertices vi and vj, (or if multiple loops are present at the same vertex), the (i, j)th entry equals the number of edges connecting the pair of vertices.
- Example: We give the adjacency matrix of the pseudograph shown here using the ordering of vertices a, b, c, d.



$$\begin{bmatrix} 0 & 3 & 0 & 2 \\ 3 & 0 & 1 & 1 \\ 0 & 1 & 1 & 2 \\ 2 & 1 & 2 & 0 \end{bmatrix}$$

Adjacency Matrices (*continued*)

- Adjacency matrices can also be used to represent directed graphs. The matrix for a directed graph $G = (V, E)$ has a 1 in its (i, j)th position if there is an edge from v_i to v_j, where $v_1, v_2, \ldots v_n$ is a list of the vertices.
  - In other words, if the graphs adjacency matrix is $\mathbf{A}_G = [a_{ij}]$, then

$$a_{ij} = \begin{cases} 1 & \text{if } \{v_i, v_j\} \text{ is an edge of } G, \\ 0 & \text{otherwise.} \end{cases}$$

  - The adjacency matrix for a directed graph does not have to be symmetric, because there may not be an edge from v_i to v_j, when there is an edge from v_j to v_i.
  - To represent directed multigraphs, the value of a_{ij} is the number of edges connecting v_i to v_j.

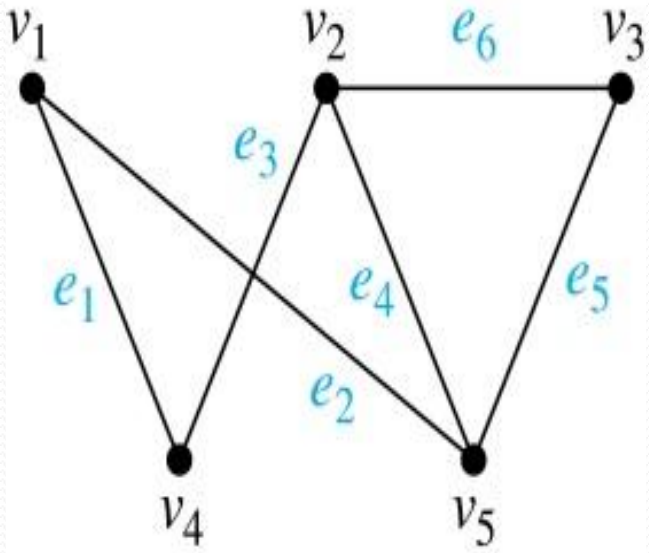
Representation of Graphs: Incidence Matrices

Definition: Let $G = (V, E)$ be an undirected graph with vertices where $v_1, v_2, \ldots v_n$ and edges $e_1, e_2, \ldots e_m$. The incidence matrix with respect to the ordering of V and E is the $n \times m$ matrix $\mathbf{M} = [m_{ij}]$, where

$$m_{ij} = \begin{cases} 1 & \text{when edge } e_j \text{ is incident with } v_i, \\ 0 & \text{otherwise.} \end{cases}$$

Incidence Matrices (*continued*)

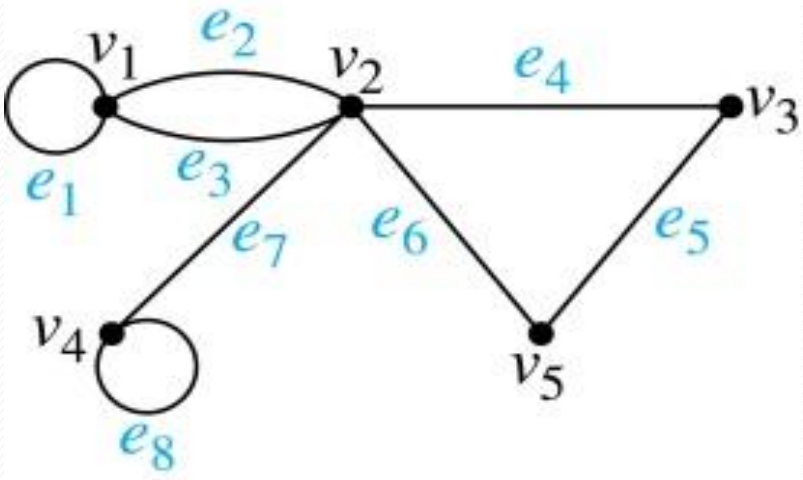
Example: Simple Graph and Incidence Matrix



$$\begin{bmatrix} 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 1 & 0 \end{bmatrix}$$

The rows going from top to bottom represent v_1 through v_5 and the columns going from left to right represent e_1 through e_6.

Example: Pseudograph and Incidence Matrix



$$\begin{bmatrix} 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 \end{bmatrix}$$

The rows going from top to bottom represent v_1 through v_5 and the columns going from left to right represent e_1 through e_8.

Connectivity

Section 10.4

Section Summary

- Paths
- Connectedness in Undirected Graphs
- Connectedness in Directed Graphs
- Counting Paths between Vertices

Paths

Informal Definition: A *path* is a sequence of edges that begins at a vertex of a graph and travels from vertex to vertex along edges of the graph. As the path travels along its edges, it visits the vertices along this path, that is, the endpoints of these.

Applications: Numerous problems can be modeled with paths formed by traveling along edges of graphs such as:

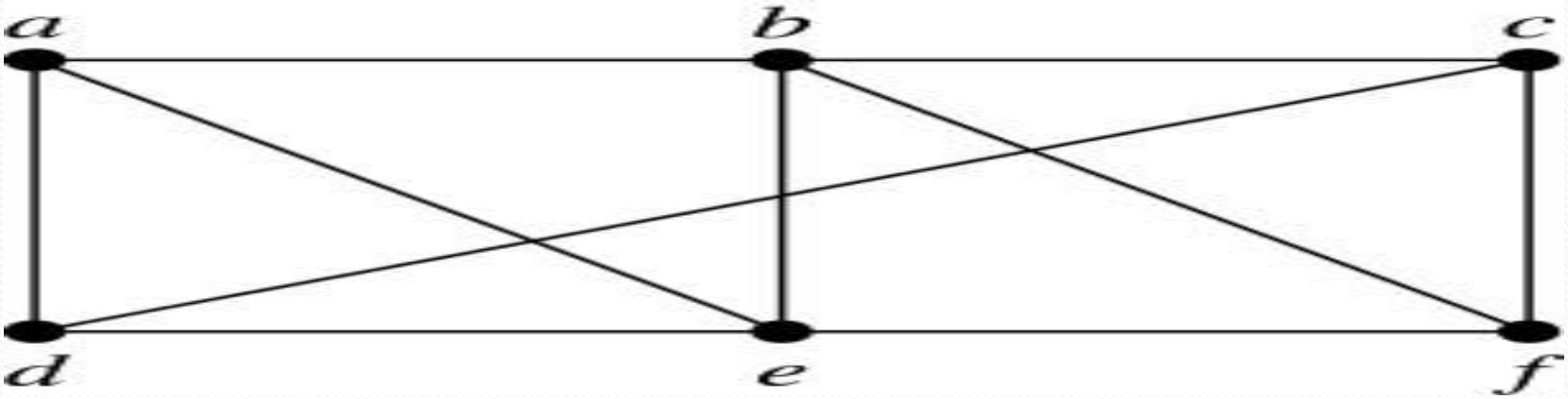
- determining whether a message can be sent between two computers.
- efficiently planning routes for mail delivery.

Paths

Definition: Let n be a nonnegative integer and G an undirected graph. A *path* of *length* n from u to v in G is a sequence of n edges $e_1, \ldots, e_n$ of G for which there exists a sequence $x_0 = u, x_1, \ldots, x_{n-1}, x_n = v$ of vertices such that e_i has, for $i = 1, \ldots, n$, the endpoints x_{i-1} and x_i.

- When the graph is simple, we denote this path by its vertex sequence $x_0, x_1, \ldots, x_n$ (since listing the vertices uniquely determines the path).
- The path is a *circuit* if it begins and ends at the same vertex ($u = v$) and has length greater than zero.
- The path or circuit is said to *pass through* the vertices $x_1, x_2, \ldots, x_{n-1}$ and *traverse* the edges $e_1, \ldots, e_n$.
- A path or circuit is *simple* if it does not contain the same edge more than once.

Paths (*continued*)



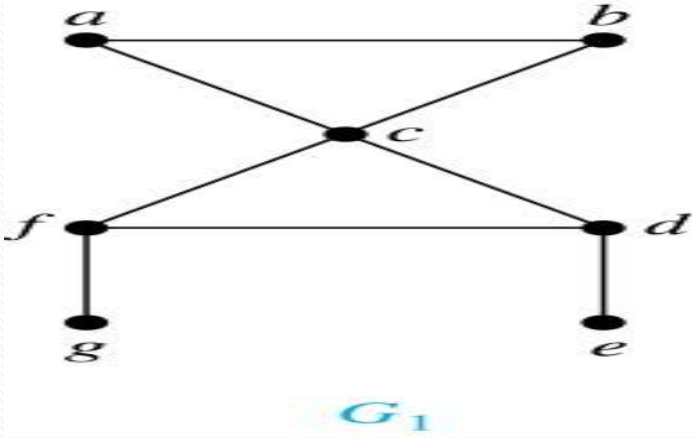
Example: In the simple graph here:

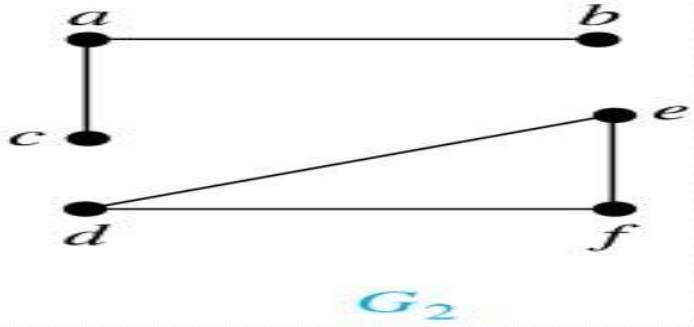
- *a*, *d*, *c*, *f*, *e* is a simple path of length 4.
- *d*, *e*, *c*, *a* is not a path because *e* is not connected to *c*.
- *b*, *c*, *f*, *e*, *b* is a circuit of length 4.
- *a*, *b*, *e*, *d*, *a*, *b* is a path of length 5, but it is not a simple path.

Connectedness in Undirected Graphs

Definition: An undirected graph is called *connected* if there is a path between every pair of vertices. An undirected graph that is not *connected* is called *disconnected*. We say that we *disconnect* a graph when we remove vertices or edges, or both, to produce a disconnected subgraph.

Example: G_1 is connected because there is a path between any pair of its vertices, as can be easily seen. However G_2 is not connected because there is no path between vertices *a* and *f*, for example.





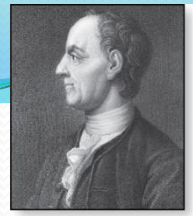
Euler and Hamiltonian Graphs

Section 10.5

Section Summary

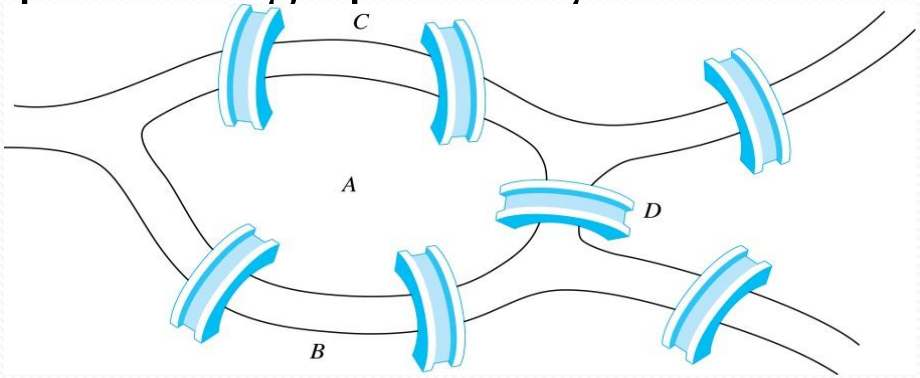
- Euler Paths and Circuits
- Hamilton Paths and Circuits
- Applications of Hamilton Circuits

Euler Paths and Circuits

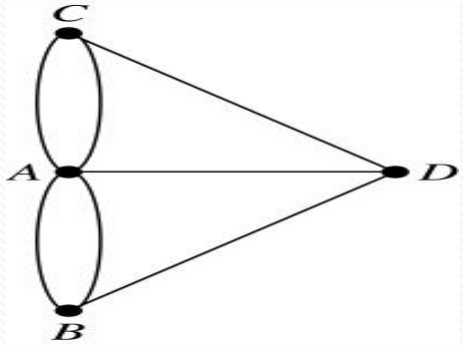


Leonard Euler
(1707-1783)

- The town of Königsberg, Prussia (now Kalingrad, Russia) was divided into four sections by the branches of the Pregel river. In the 18th century seven bridges connected these regions.
- People wondered whether whether it was possible to follow a path that crosses each bridge exactly once and returns to the starting point.
- The Swiss mathematician Leonard Euler proved that no such path exists. This result is often considered to be the first theorem ever proved in graph theory.



The 7 Bridges of Königsberg

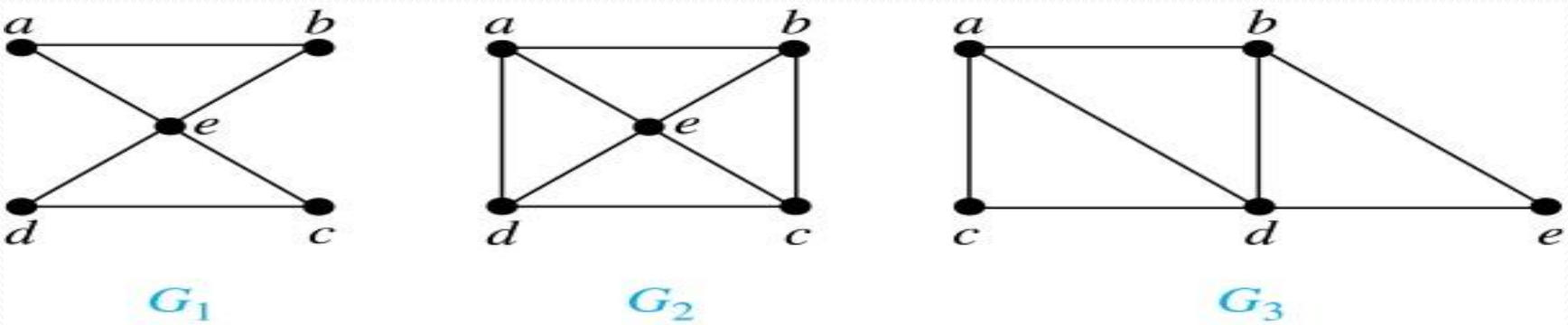


Multigraph Model of the Bridges of Königsberg

Euler Paths and Circuits

Definition: An *Euler circuit* in a graph G is a simple circuit containing every edge of G. An *Euler path* in G is a simple path containing every edge of G.

Example: Which of the undirected graphs G_1, G_2, and G_3 has a Euler circuit? Of those that do not, which has an Euler path?



Solution: The graph G_1 has an Euler circuit (e.g., a, e, c, d, e, b, a). But, as can easily be verified by inspection, neither G_2 nor G_3 has an Euler circuit. Note that G_3 has an Euler path (e.g., a, c, d, e, b, d, a, b), but there is no Euler path in G_2, which can be verified by inspection.

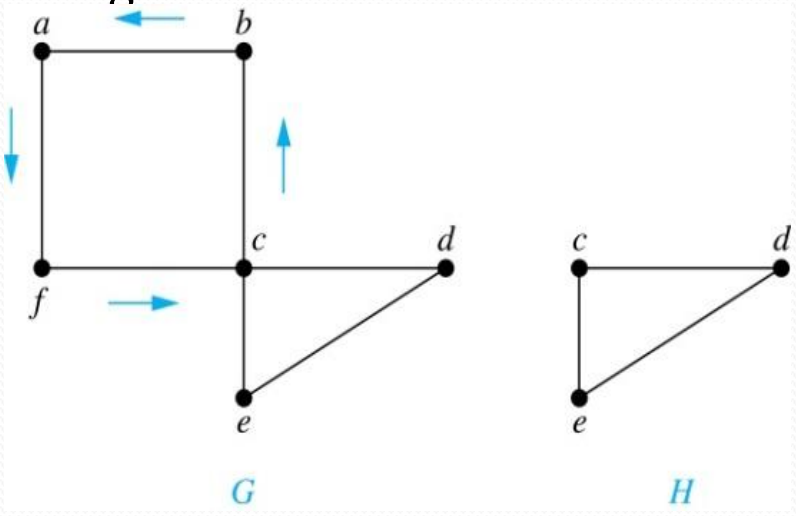
Necessary Conditions for Euler Circuits and Paths

- An Euler circuit begins with a vertex a and continues with an edge incident with a, say $\{a, b\}$. The edge $\{a, b\}$ contributes one to $\deg(a)$.
- Each time the circuit passes through a vertex it contributes two to the vertex's degree.
- Finally, the circuit terminates where it started, contributing one to $\deg(a)$. Therefore $\deg(a)$ must be even.
- We conclude that the degree of every other vertex must also be even.
- By the same reasoning, we see that the initial vertex and the final vertex of an Euler path have odd degree, while every other vertex has even degree. So, a graph with an Euler path has exactly two vertices of odd degree.
- In the next slide we will show that these necessary conditions are also sufficient conditions.

Sufficient Conditions for Euler Circuits and Paths

Suppose that G is a connected multigraph with ≥ 2 vertices, all of even degree. Let $x_0 = a$ be a vertex of even degree. Choose an edge $\{x_0, x_1\}$ incident with a and proceed to build a simple path $\{x_0, x_1\}, \{x_1, x_2\}, \ldots, \{x_{n-1}, x_n\}$ by adding edges one by one until another edge can not be added.

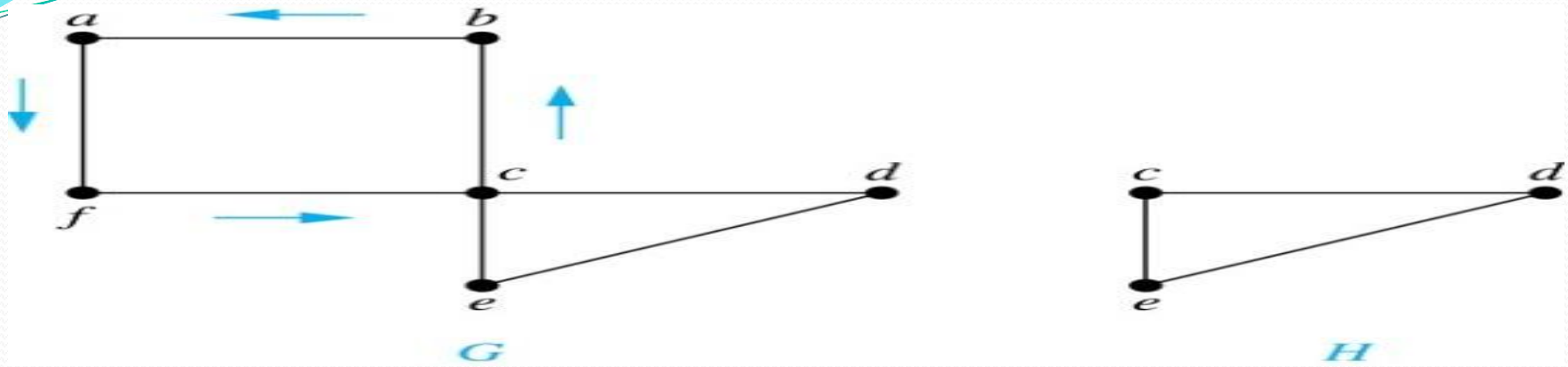
We illustrate this idea in the graph G here. We begin at a and choose the edges $\{a, f\}, \{f, c\}, \{c, b\}$, and $\{b, a\}$ in succession.



- The path begins at a with an edge of the form $\{a, x\}$; we show that it must terminate at a with an edge of the form $\{y, a\}$. Since each vertex has an even degree, there must be an even number of edges incident with this vertex. Hence, every time we enter a vertex other than a, we can leave it. Therefore, the path can only end at a.
- If all of the edges have been used, an Euler circuit has been constructed. Otherwise, consider the subgraph H obtained from G by deleting the edges already used.

In the example H consists of the vertices c, d, e.

Sufficient Conditions for Euler Circuits and Paths



- Because G is connected, H must have at least one vertex in common with the circuit that has been deleted.

In the example, the vertex is *c*.

- Every vertex in H must have even degree because all the vertices in *G* have even degree and for each vertex, pairs of edges incident with this vertex have been deleted. Beginning with the shared vertex construct a path ending in the same vertex (as was done before). Then splice this new circuit into the original circuit.

In the example, we end up with the circuit *a, f, c, d, e, c, b, a*.

- Continue this process until all edges have been used. This produces an Euler circuit. Since every edge is included and no edge is included more than once.
- Similar reasoning can be used to show that a graph with exactly two vertices of odd degree must have an Euler path connecting these two vertices of odd degree

Algorithm for Constructing an Euler Circuits

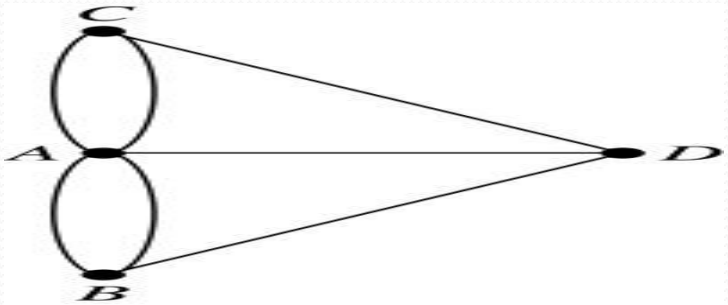
In our proof we developed this algorithms for constructing a Euler circuit in a graph with no vertices of odd degree.

procedure *Euler*(*G*: connected multigraph with all vertices of even degree)
circuit := a circuit in *G* beginning at an arbitrarily chosen vertex with edges successively added to form a path that returns to this vertex.
H := *G* with the edges of this circuit removed
while *H* has edges
subciruit := a circuit in *H* beginning at a vertex in *H* that also is an endpoint of an edge in circuit.
H := *H* with edges of *subciruit* and all isolated vertices removed
circuit := *circuit* with *subcircuit* inserted at the appropriate vertex.
return *circuit*{*circuit* is an Euler circuit}

Necessary and Sufficient Conditions for Euler Circuits and Paths (*continued*)

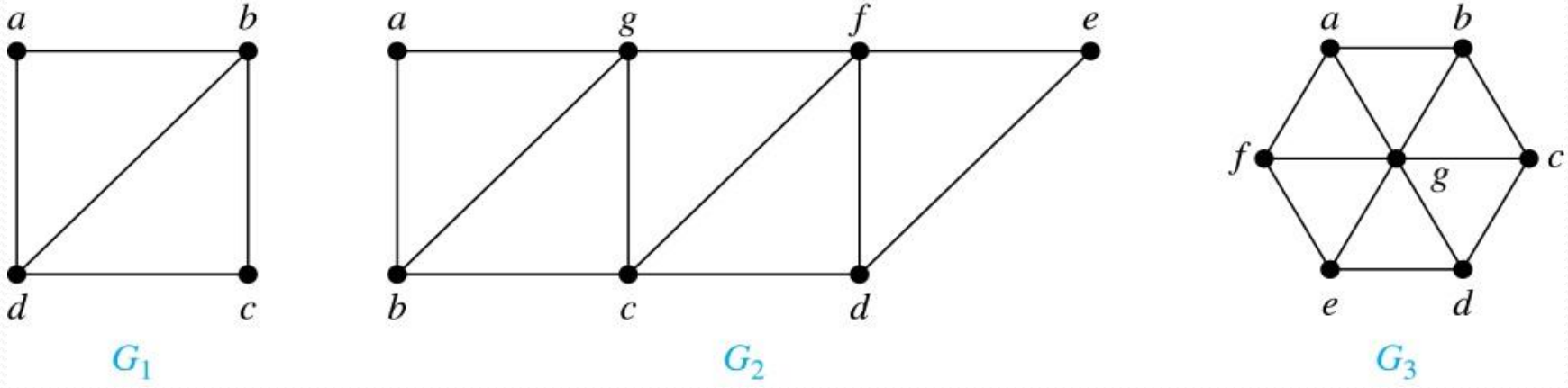
Theorem: A connected multigraph with at least two vertices has an Euler circuit if and only if each of its vertices has an even degree and it has an Euler path if and only if it has exactly two vertices of odd degree.

Example: Two of the vertices in the multigraph model of the Königsberg bridge problem have odd degree. Hence, there is no Euler circuit in this multigraph and it is impossible to start at a given point, cross each bridge exactly once, and return to the starting point.



Euler Circuits and Paths

Example:



G_1 contains exactly two vertices of odd degree (b and d). Hence it has an Euler path, e.g., d, a, b, c, d, b.

G_2 has exactly two vertices of odd degree (b and d). Hence it has an Euler path, e.g., $b, a, g, f, e, d, c, g, b, c, f, d$.

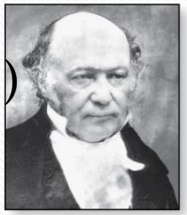
G_3 has six vertices of odd degree. Hence, it does not have an Euler path.

Applications of Euler Paths and Circuits

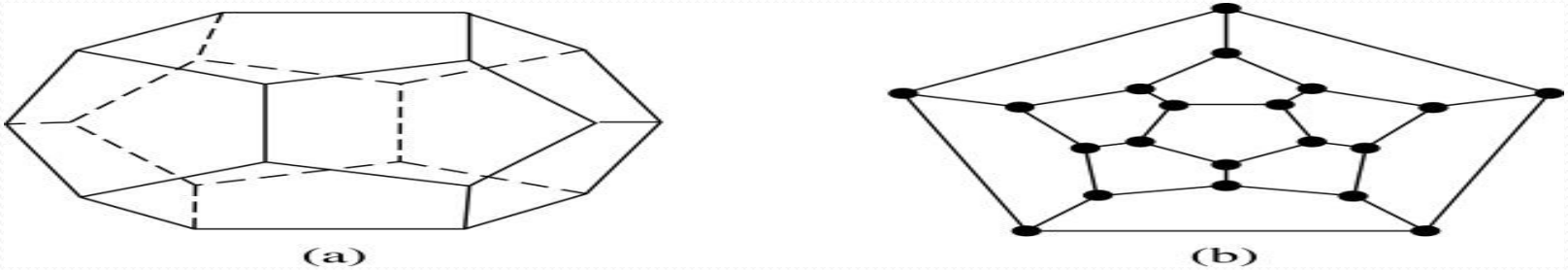
- Euler paths and circuits can be used to solve many practical problems such as finding a path or circuit that traverses each
  - street in a neighborhood,
  - road in a transportation network,
  - connection in a utility grid,
  - link in a communications network.
- Other applications are found in the
  - layout of circuits,
  - network multicasting,
  - molecular biology, where Euler paths are used in the sequencing of DNA.

Hamilton Paths and Circuits

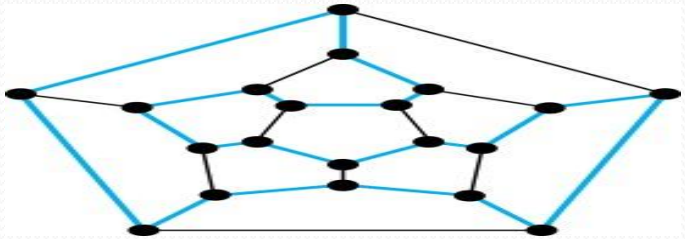
William Rowan Hamilton (1805- 1865)



- Euler paths and circuits contained every edge only once.
  Now we look at paths and circuits that contain every vertex exactly once.
- William Hamilton invented the *Icosian puzzle* in 1857. It consisted of a wooden dodecahedron (with 12 regular pentagons as faces), illustrated in (a), with a peg at each vertex, labeled with the names of different cities. String was used to used to plot a circuit visiting 20 cities exactly once
- The graph form of the puzzle is given in (b).



- The solution (a Hamilton circuit) is given here.



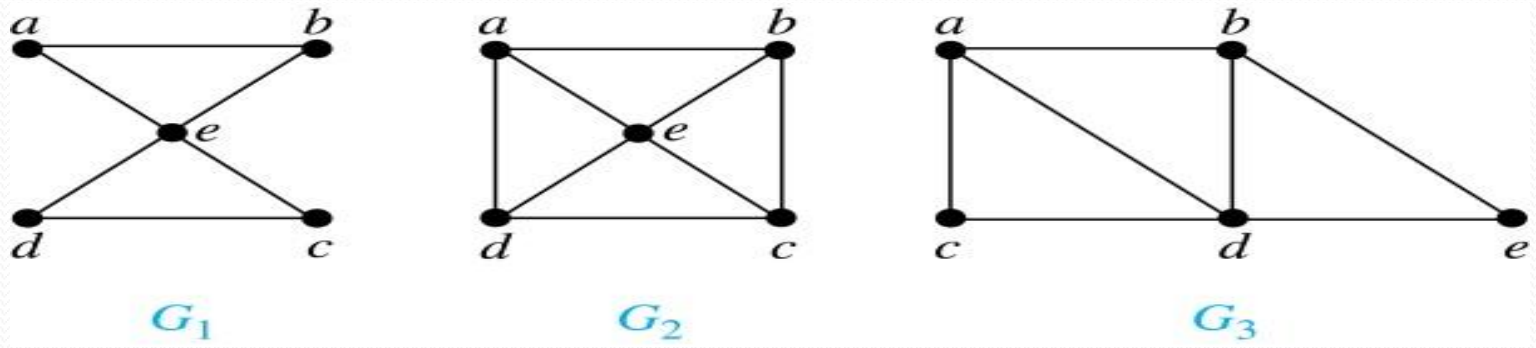
Hamilton Paths and Circuits

Definition: A simple path in a graph G that passes through every vertex exactly once is called a *Hamilton path*, and a simple circuit in a graph G that passes through every vertex exactly once is called a *Hamilton circuit*.

That is, a simple path $x_0, x_1, \ldots, x_{n-1}, x_n$ in the graph $G = (V, E)$ is called a Hamilton path if $V = \{x_0, x_1, \ldots, x_{n-1}, x_n\}$ and $x_i \neq x_j$ for $0 \leq i < j \leq n$, and the simple circuit $x_0, x_1, \ldots, x_{n-1}, x_n, x_0$ (with $n > 0$) is a Hamilton circuit if $x_0, x_1, \ldots, x_{n-1}, x_n$ is a Hamilton path.

Hamilton Paths and Circuits

Example: Which of these simple graphs has a Hamilton circuit or, if not, a Hamilton path?



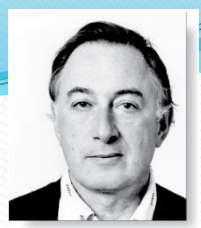
Solution:

G_1 does not have a Hamilton circuit (Why?), but does have a Hamilton path : a, b ,e , c, d.

G_2 has a Hamilton circuit: a, b, c, d, e, a.

G_3 has a Hamilton circuit: a,b,e,d,c,a

Necessary Conditions for Hamilton Circuits



Gabriel Andrew Dirac
(1925-1984)

- Unlike for an Euler circuit, no simple necessary and sufficient conditions are known for the existence of a Hamiton circuit.
- However, there are some useful necessary conditions. We describe two of these now.

  Dirac's Theorem: If G is a simple graph with $n \geq 3$ vertices such that the degree of every vertex in G is $\geq n/2$, then G has a Hamilton circuit.

  Ore's Theorem: If G is a simple graph with $n \geq 3$ vertices such that $\deg(u) + \deg(v) \geq n$ for every pair of nonadjacent vertices, then G has a Hamilton circuit.

Øysten Ore
(1899-1968)

Applications of Hamilton Paths and Circuits

- Applications that ask for a path or a circuit that visits each intersection of a city, each place pipelines intersect in a utility grid, or each node in a communications network exactly once, can be solved by finding a Hamilton path in the appropriate graph.
- The famous *traveling salesperson problem* (*TSP*) asks for the shortest route a traveling salesperson should take to visit a set of cities. This problem reduces to finding a Hamilton circuit such that the total sum of the weights of its edges is as small as possible.
- A family of binary codes, known as *Gray codes*, which minimize the effect of transmission errors, correspond to Hamilton circuits in the n-cube Q_n.

ISOMORPHISM OF GRAPHS

Section 10.3

Isomorphism of Graphs

Definition: The simple graphs $G_1 = (V_1, E_1)$ and $G_2 = (V_2, E_2)$ are *isomorphic* if there is a one-to-one and onto function f from V_1 to V_2 with the property that a and b are adjacent in G_1 if and only if $f(a)$ and $f(b)$ are adjacent in G_2 , for all a and b in V_1 . Such a function f is called an *isomorphism*. Two simple graphs that are not isomorphic are called *nonisomorphic*.

Isomorphism of Graphs

- It is difficult to determine whether two simple graphs are isomorphic using brute force because there are $n!$ possible one-to-one correspondences between the vertex sets of two simple graphs with n vertices.
- The best algorithms for determining weather two graphs are isomorphic have exponential worst case complexity in terms of the number of vertices of the graphs.
- Sometimes it is not hard to show that two graphs are not isomorphic. We can do so by finding a property, preserved by isomorphism, that only one of the two graphs has. Such a property is called *graph invariant*.
- There are many different useful graph invariants that can be used to distinguish nonisomorphic graphs, such as the number of vertices, number of edges, and degree sequence (list of the degrees of the vertices in nonincreasing order). We will encounter others in later sections of this chapter.

ISOMORPHIC INVARIANT

- A property P is called an isomorphic invariant if, and only if, given any graphs G and G',
- If G has property P and G' is isomorphic to G, then G' has property P.

THEOREM OF ISOMORPHIC INVARIANT

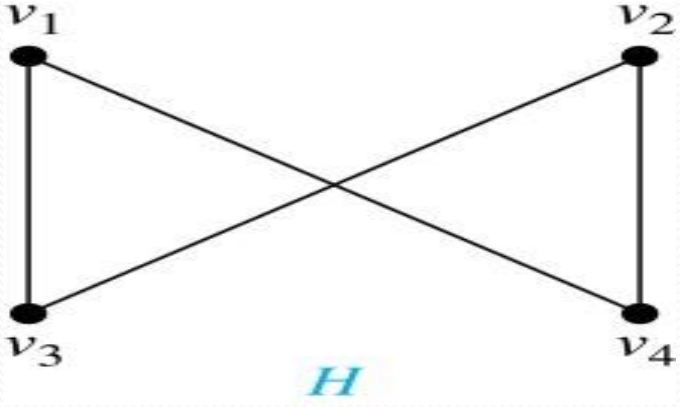
Each of the following properties is an invariant for graph isomorphism, where n, m and k are all non-negative integers, if the graph:

1. has n vertices.
2. has m edges.
3. has a vertex of degree k.
4. has m vertices of degree k.
5. has a circuit of length k.
6. has a simple circuit of length k.
7. has m simple circuits of length k.
8. is connected.
9. has an Euler circuit.
10. has a Hamiltonian circuit.

Isomorphism of Graphs (*cont.*)

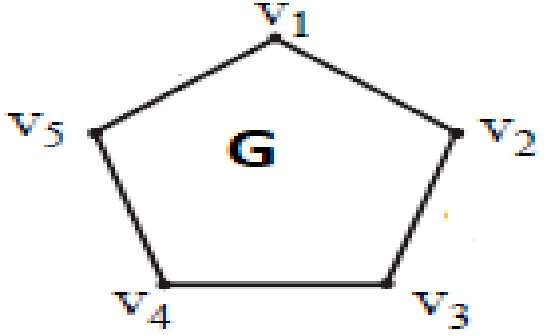
Example: Show that the graphs $G = (V, E)$ and $H = (W, F)$ are isomorphic.

Solution: The function f with $f(u_1) = v_1, f(u_2) = v_4, f(u_3) = v_3$, and $f(u_4)$ correspondence between V and W. Note that adjacent vertices in G are u_1 and u_2, u_1 and u_3, u_2 and u_4, and u_3 and u_4. Each of the pairs $f(u_1) = v_1$ and $f(u_2) = v_4$, $f(u_1) = v_1$ and $f(u_3) = v_3$, $f(u_2) = v_4$ and $f(u_4) = v_2$, and $f(u_3) = v_3$ and $f(u_4) = v_2$ consists of two adjacent vertices in H.



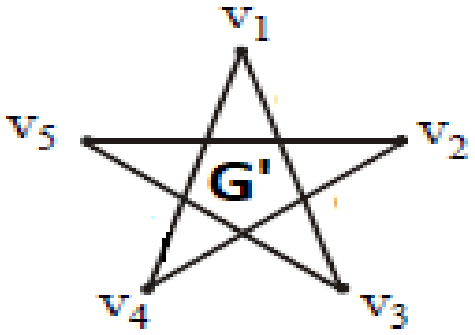
Isomorphism of Graphs

Example: Show that the graphs $G = (V, E)$ and $G' = (W, F)$ are isomorphic.



Solution: The function f with

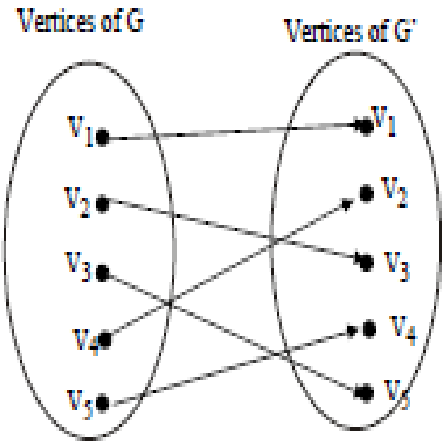
f(v1) = v1,f(v2) = v3, f(v3) = v5, f(v4) = v2 and f(v5) = v4 is a one-to-one correspondence between V and W.



Note that adjacent vertices in G are v1 and v2, v2 and v3, v3 and v4, v4 and v5 and v5 and v1.

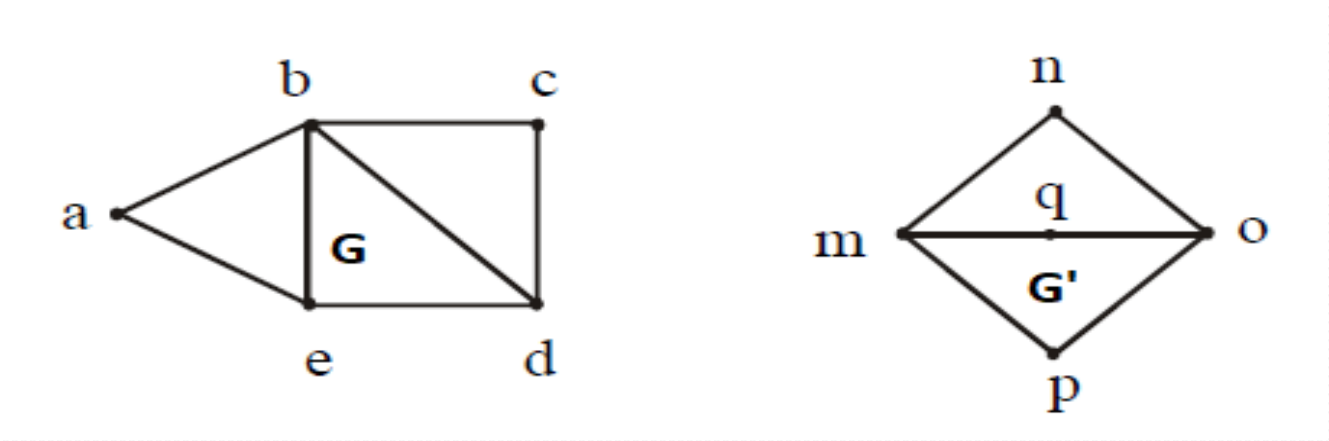
Each of the pairs f(v1) = v1 and f(v2) = v3, f(v2) = v3 and f(v3) = v5 , f(v3) = v5 and

f(v4) = v2 , f(v4) = v2 and f(v5) = v4 and f(v5) = v4 and f(v1) = v1 consists of two adjacent vertices in H.



Isomorphism of Graphs

EXAMPLE: Determine whether the graph G and G' given below are isomorphic.



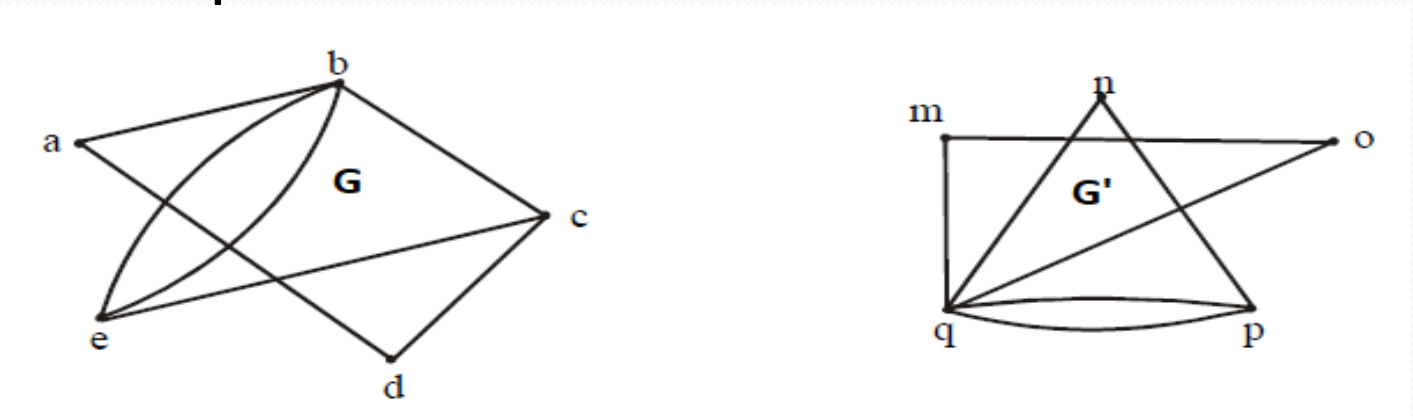
SOLUTION:

As both the graphs have the same number of vertices. But the graph G has 7 edges and the graph G' has only 6 edges. Therefore the two graphs are not isomorphic.

Note: As the edges of both the graphs G and G' are not same then how the one-one correspondence is possible ,that the reason the graphs G and G' are not isomorphic.

Isomorphism of Graphs

EXAMPLE: Determine whether the graph G and G' given below are isomorphic.

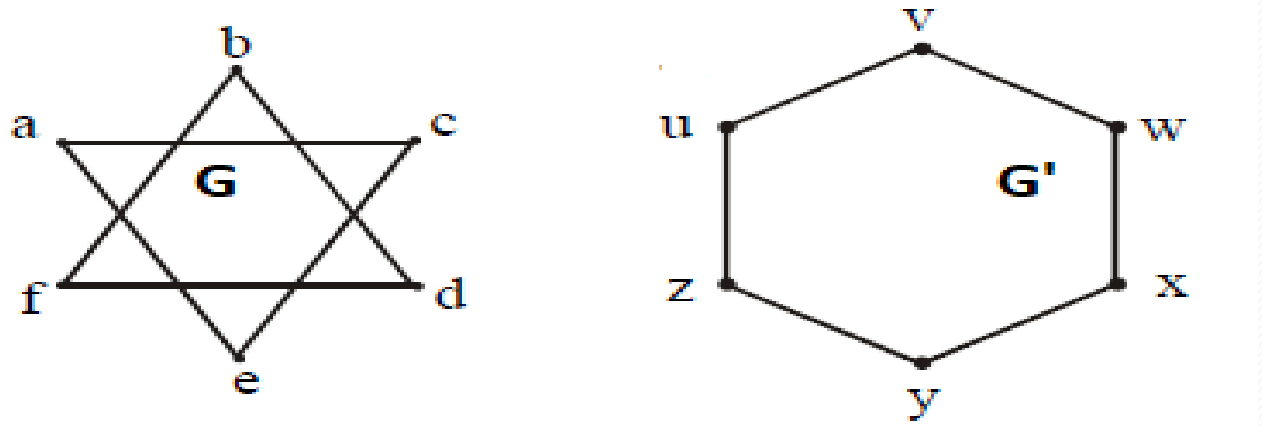


SOLUTION:

Both the graphs have 5 vertices and 7 edges. The vertex q of G' has degree 5. However G does not have any vertex of degree 5 (so one-one correspondence is not possible). Hence, the two graphs are not isomorphic.

Isomorphism of Graphs

EXAMPLE: Determine whether the graph G and G' given below are isomorphic.

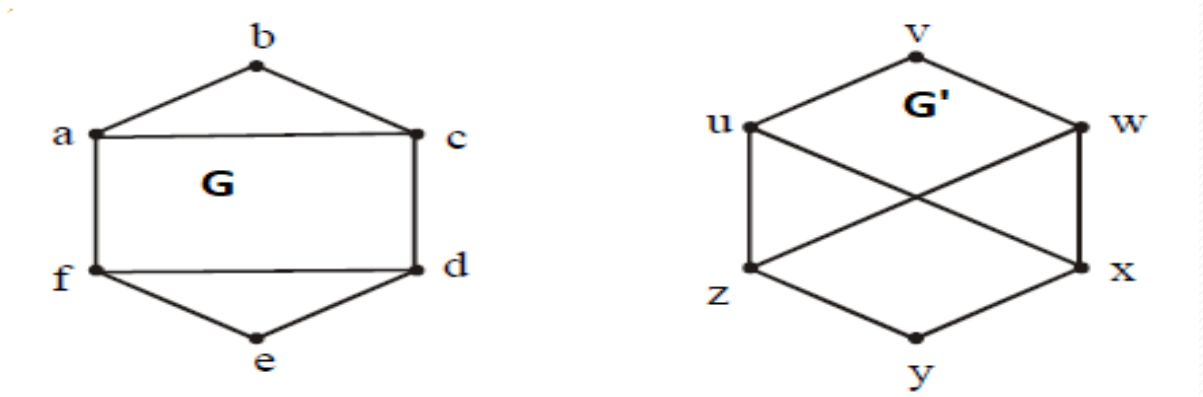


SOLUTION:

Clearly the vertices of both the graphs G and G' have the same degree (i.e "2") and having the same number of vertices and edges but isomorphism is not possible. As the graph G' is a connected graph but the graph G is not connected due to have two components (eca and bdf). Therefore the two graphs are non isomorphic.

Isomorphism of Graphs

EXAMPLE: Determine whether the graph G and G' given below are isomorphic.



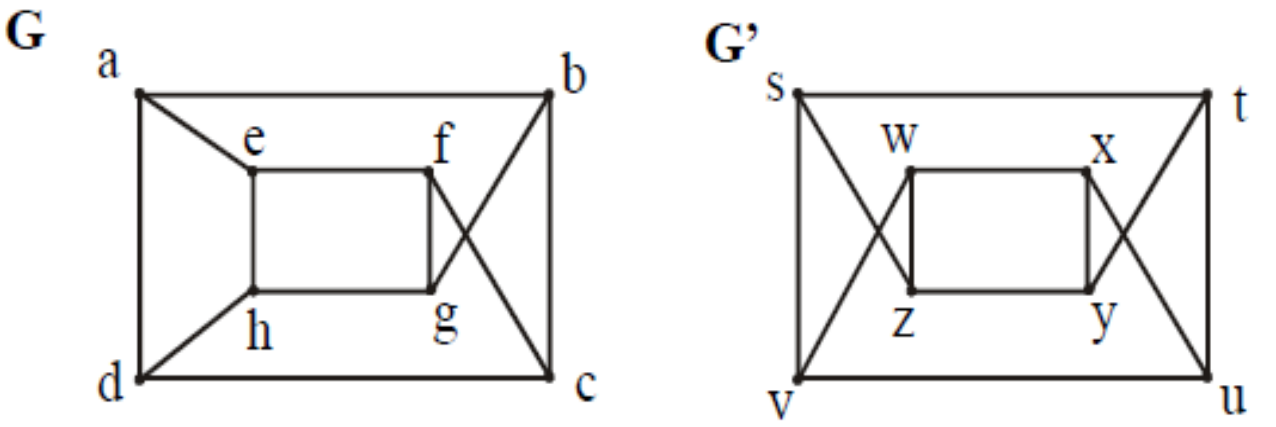
SOLUTION:

Clearly G has six vertices, G' also has six vertices. And the graph G has two simple circuits of length 3; one is abca and the other is defd. But G' does not have any simple circuit of length 3(as one simple circuit in G' is uxwv of length 4).Therefore the two graphs are non-isomorphic.

Note: A simple circuit is a circuit that does not have any other repeated vertex except the first and last.

Isomorphism of Graphs

EXAMPLE: Determine whether the graph G and G' given below are isomorphic.



SOLUTION:

Both the graph G and G' have 8 vertices and 12 edges and both are also called regular graph(as each vertex has degree 3).The graph G has two simple circuits of length 5; abcfea(i.e starts and ends at a) and cdhgfc(i.e starts and ends at c). But G' does not have any simple circuit of length 5 (it has simple circuit tyxut,vwxuv of length 4 etc). Therefore the two graphs are non-isomorphic.

Isomorphism of Graphs

EXAMPLE: Determine whether the given graph G and H are isomorphic.

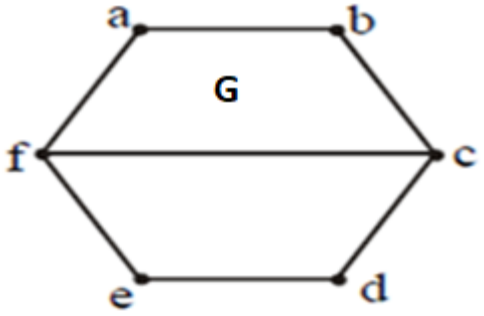
Solution:

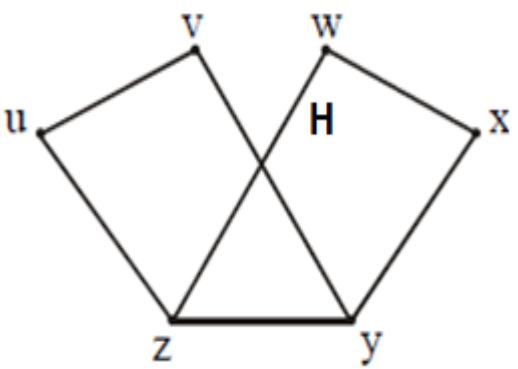
Solution: The function f with

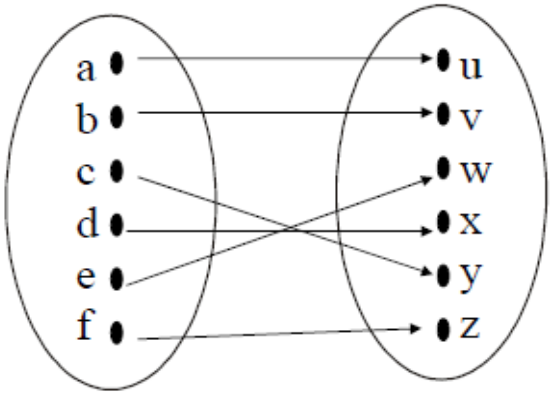
f(a) = u, f(b) = v, f(c) = y, f(d) = x, f(e) = w and f(f) = z is a one-to-one correspondence between G and H.

Note that adjacent vertices in G are a and b, b and c, c and d, c and f, d and e, e and f and f and a.

Each of the pairs f(a) = u and f(b) = v, f(b) = v and f(c) = y , f(c) = y and f(d) = x , f(c) = y and f(f) = z , f(d) = x and f(e) = w, f(e) = w and f(f) = z and f(f) = z and f(a) = u consists of two adjacent vertices in H.

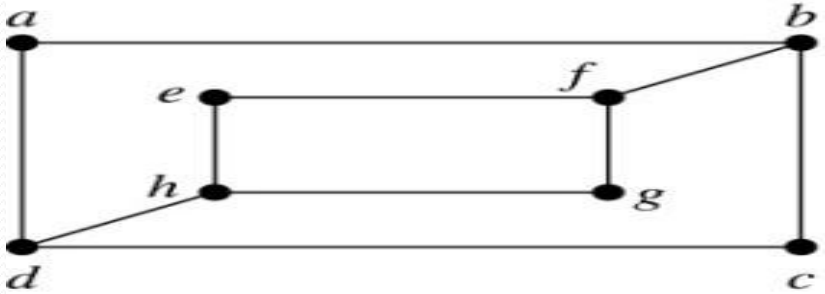




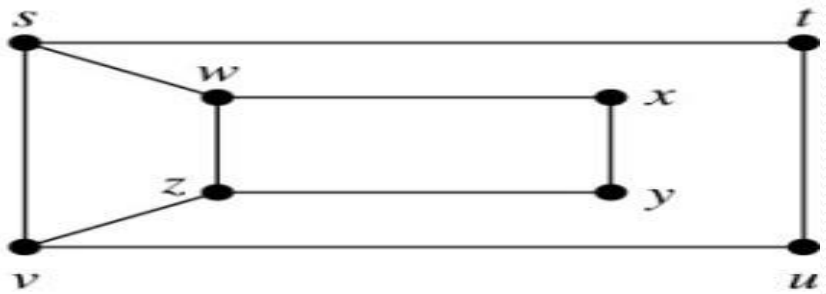


Isomorphism of Graphs (*cont.*)

Example: Determine whether these two graphs are isomorphic.



G

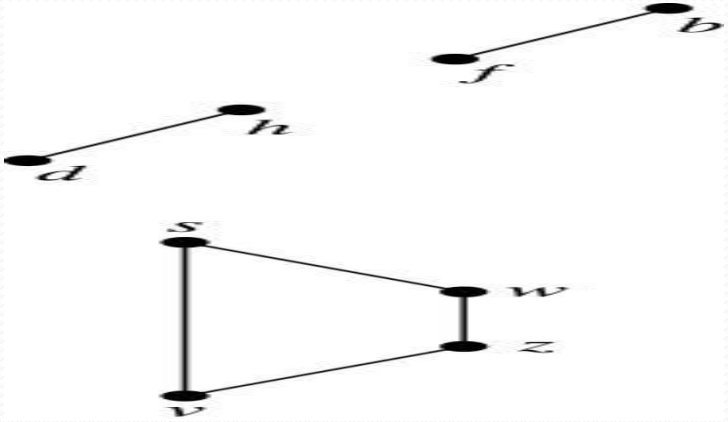


H

Solution: Both graphs have eight vertices and ten edges.
They also both have four vertices of degree two and four of degree three.

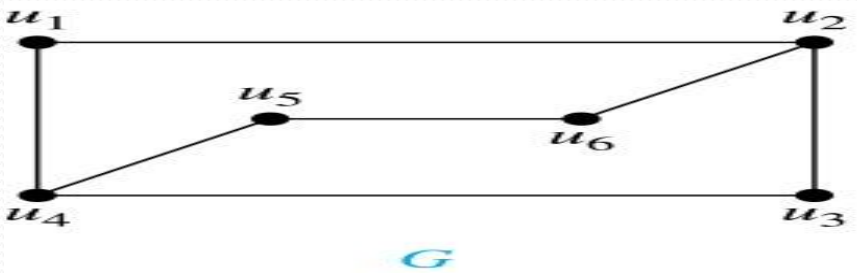
However, *G* and *H* are not isomorphic. Note that since *deg*(*a*) = 2 in *G*, *a* must correspond to *t*, *u*, *x*, or *y* in H, because these are the vertices of degree 2. But each of these vertices is adjacent to another vertex of degree two in *H*, which is not true for *a* in *G*.

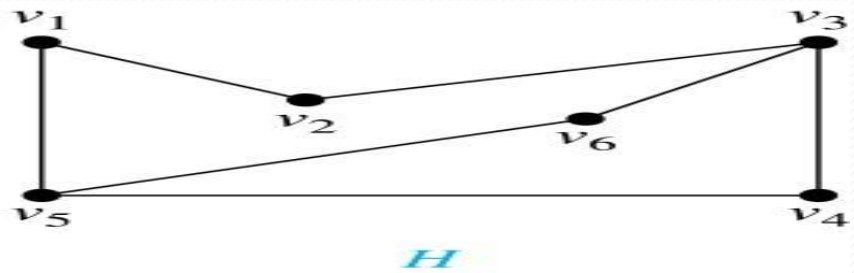
Alternatively, note that the subgraphs of *G* and *H* made up of vertices of degree three and the edges connecting them must be isomorphic.
But the subgraphs, as shown at the right, are not isomorphic.



Isomorphism of Graphs

Example: Determine whether these two graphs are isomorphic.





Solution: Both graphs have six vertices and seven edges.

They also both have four vertices of degree two and two of degree three. The subgraphs of G and H consisting of all the vertices of degree two and the edges connecting them are isomorphic. So, it is reasonable to try to find an isomorphism f.

We define an injection f from the vertices of G to the vertices of H that preserves the degree of vertices. We will determine whether it is an isomorphism.

The function f with $f(u_1) = v_6$, $f(u_2) = v_3$, $f(u_3) = v_4$, and $f(u_4) = v_5$, $f(u_5) = v_1$, and $f(u_6) = v_2$ is a one-to-one correspondence between G and H. Showing that this correspondence preserves edges is straightforward, so we will omit the details here. Because f is an isomorphism, it follows that G and H are isomorphic graphs.

Algorithms for Graph Isomorphism

- The best algorithms known for determining whether two graphs are isomorphic have exponential worst-case time complexity (in the number of vertices of the graphs).
- However, there are algorithms with linear average-case time complexity.
- You can use a public domain program called NAUTY to determine in less than a second whether two graphs with as many as 100 vertices are isomoprhic.
- Graph isomorphism is a problem of special interest because it is one of a few NP problems not known to be either tractable or NP-complete (see Section 3.3).

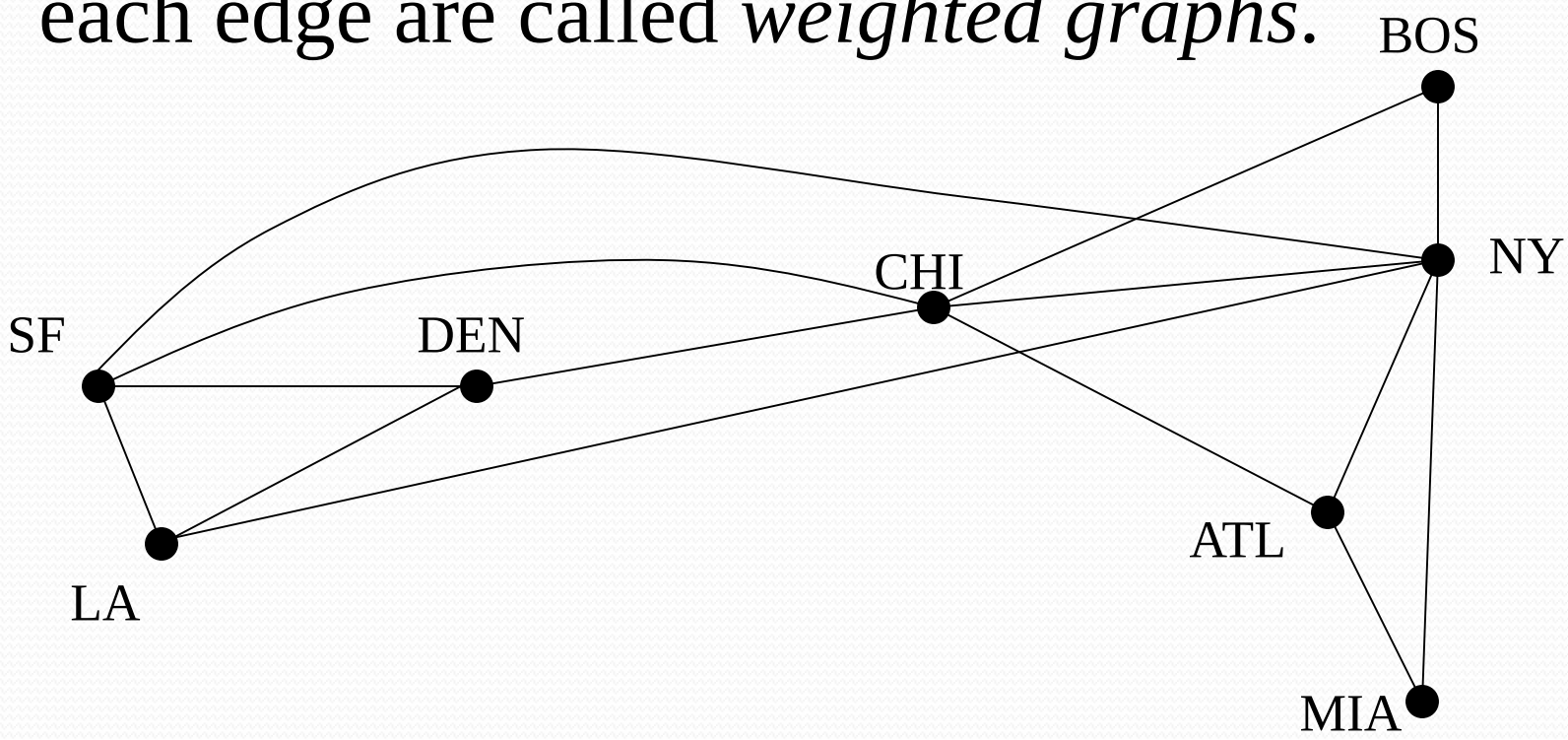
Applications of Graph Isomorphism

- The question whether graphs are isomorphic plays an important role in applications of graph theory. For example,
  - chemists use molecular graphs to model chemical compounds. Vertices represent atoms and edges represent chemical bonds. When a new compound is synthesized, a database of molecular graphs is checked to determine whether the graph representing the new compound is isomorphic to the graph of a compound that this already known.
  - Electronic circuits are modeled as graphs in which the vertices represent components and the edges represent connections between them. Graph isomorphism is the basis for
    - the verification that a particular layout of a circuit corresponds to the design's original schematics.
    - determining whether a chip from one vendor includes the intellectual property of another vendor.

Shortest Paths

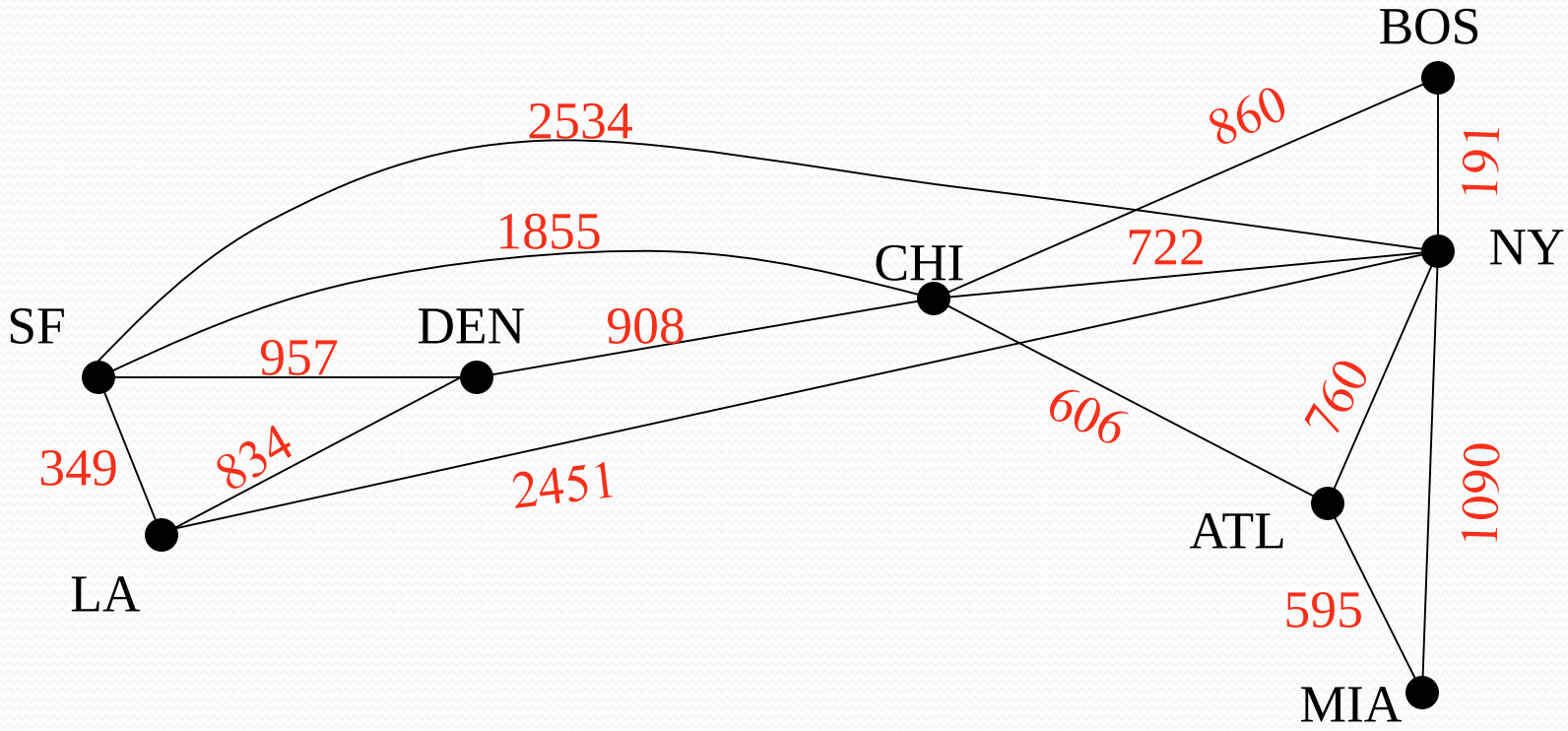
Weighted Graphs

Graphs that have a number assigned to each edge are called *weighted graphs*.



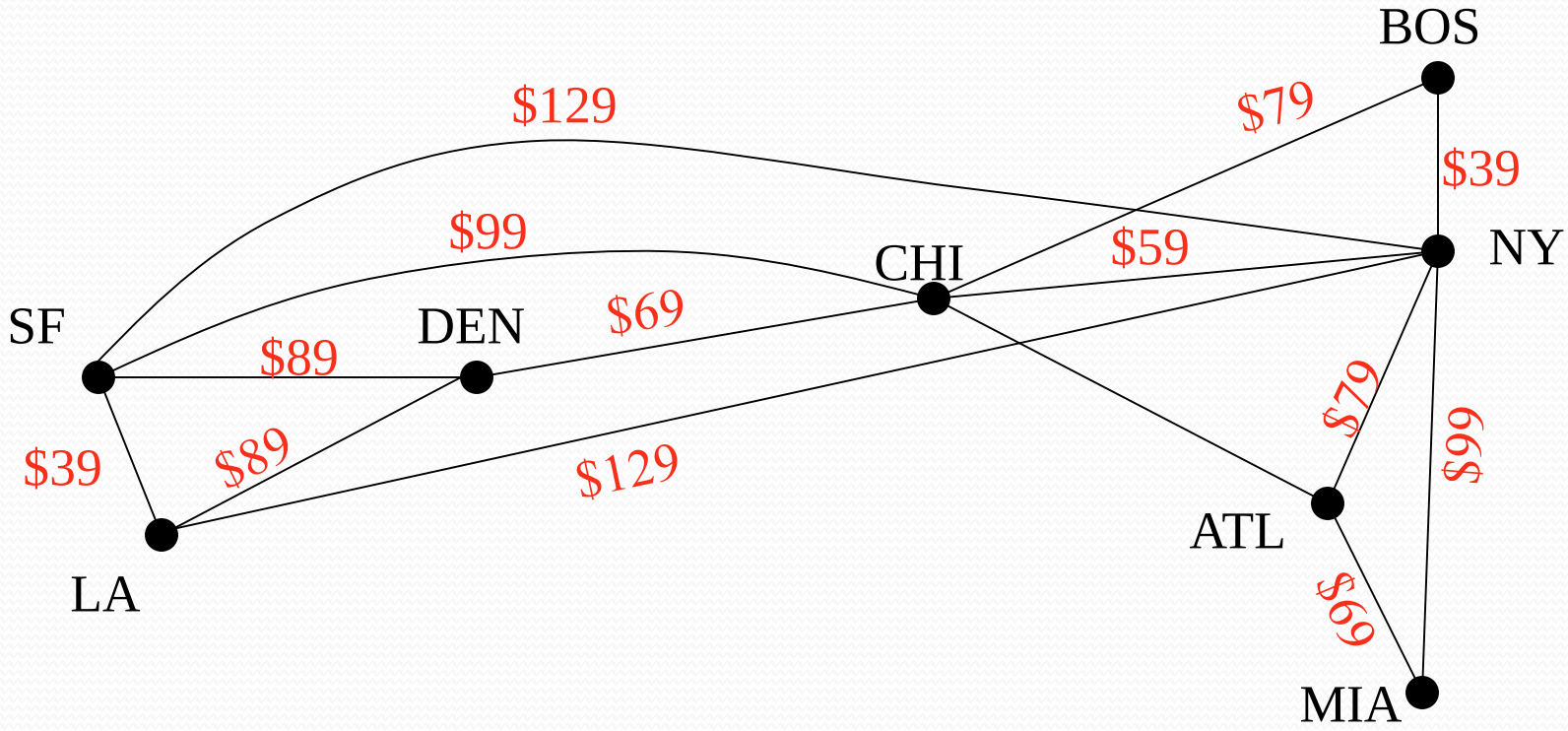
Weighted Graphs

MILES



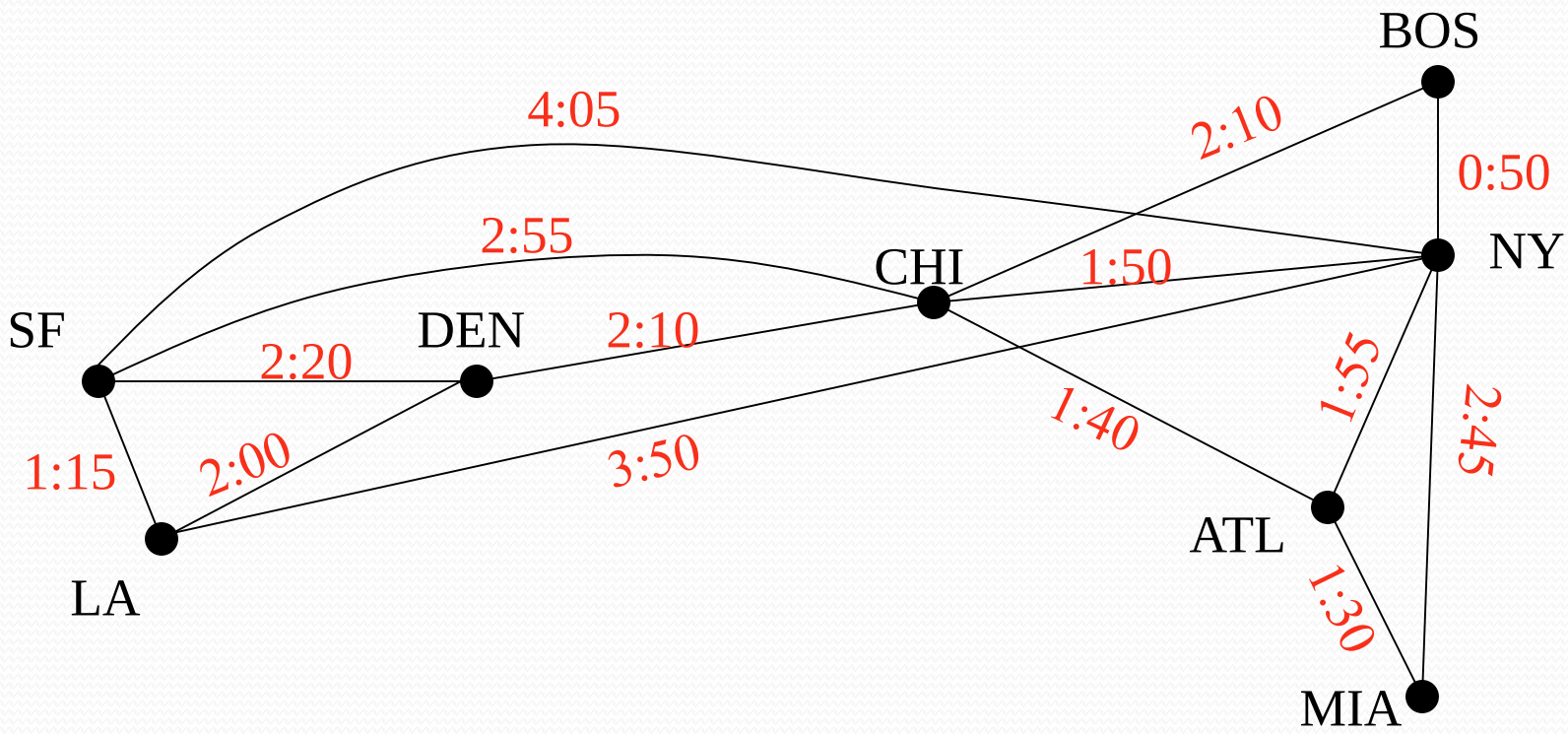
Weighted Graphs

FARES



Weighted Graphs

FLIGHT TIMES



Weighted Graphs

- A weighted graph is a graph in which each edge (u, v) has a weight w(u, v). Each weight is a real number.
- Weights can represent distance, cost, time, capacity, etc.
- The length of a path in a weighted graph is the sum of the weights on the edges.
- Dijkstra's Algorithm finds the shortest path between two vertices.

Dijkstra's Algorithm

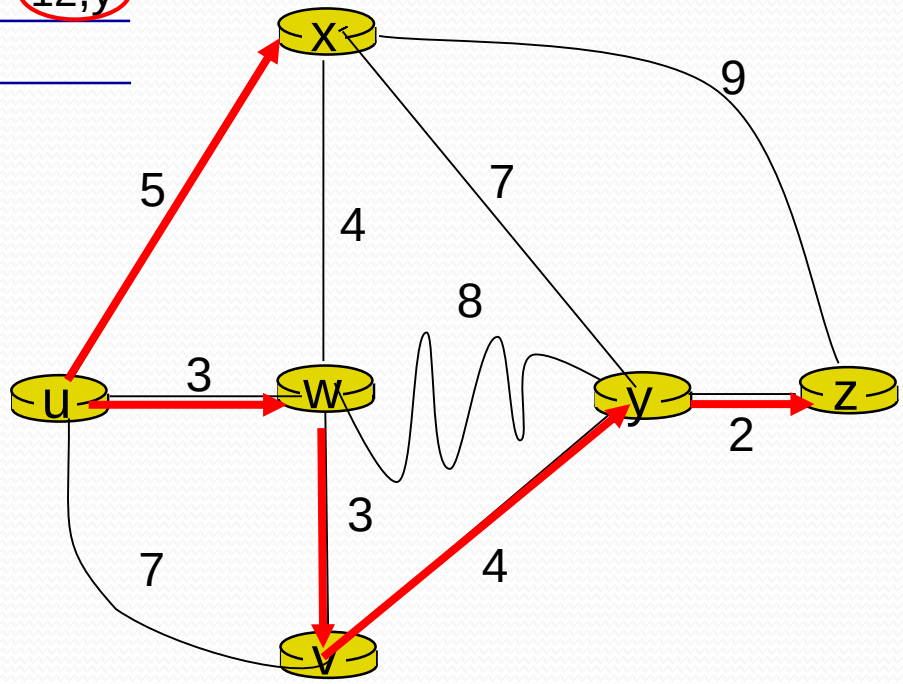
- Dijkstra's algorithm is used in problems relating to finding the shortest path.
- Each node is given a temporary label denoting the length of the shortest path *from* the start node *so far*.
- This label is replaced if another shorter route is found.
- Once it is certain that no other shorter paths can be found, the temporary label becomes a permanent label.
- Eventually all the nodes have permanent labels.
- At this point the shortest path is found by retracing the path backwards.

Dijkstra's algorithm: example

Step	N'	D(v)	D(w)	D(x)	D(y)	D(z)
0	u	7,u	3,u	5,u	∞	∞
1	uw	6,w		5,u	11,w	∞
2	uwx	6,w			11,w	14,x
3	uwxv				10,v	14,x
4	uwxvy					12,y
5	uwxvyz					

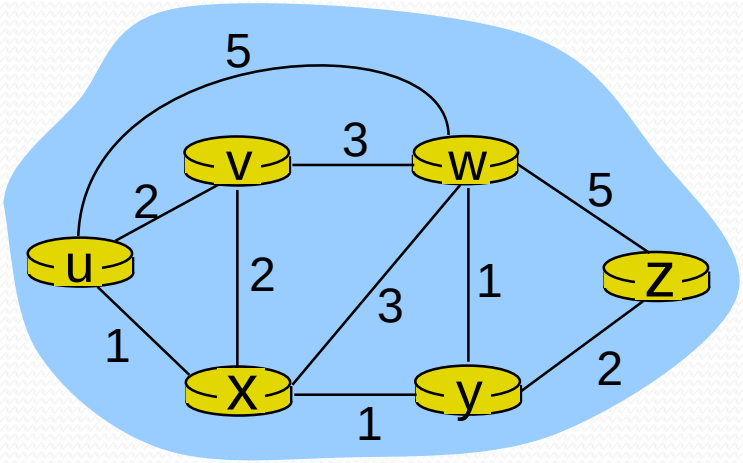
notes:

- construct shortest path tree by tracing predecessor nodes
- ties can exist (can be broken arbitrarily)



Dijkstra's algorithm: another example

Step	N'	D(v)	D(w)	D(x)	D(y)	D(z)
0	u	2,u	5,u	1,u	∞	∞
1	ux	2,u	4,x		2,x	∞
2	uxy	2,u	3,y			4,y
3	uxyv		3,y			4,y
4	uxyvw					4,y
5	uxyvwz					



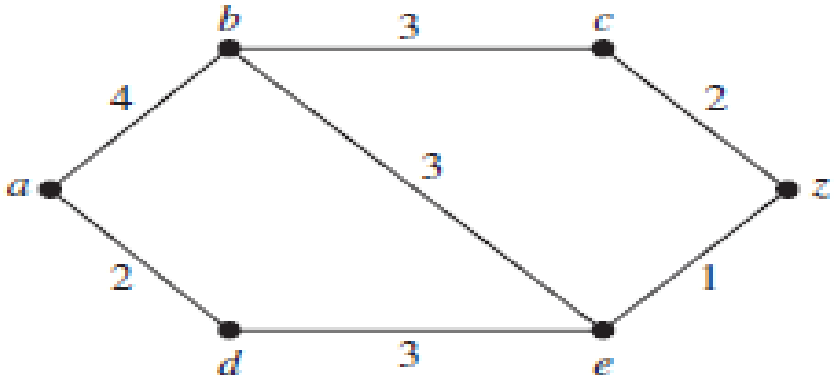
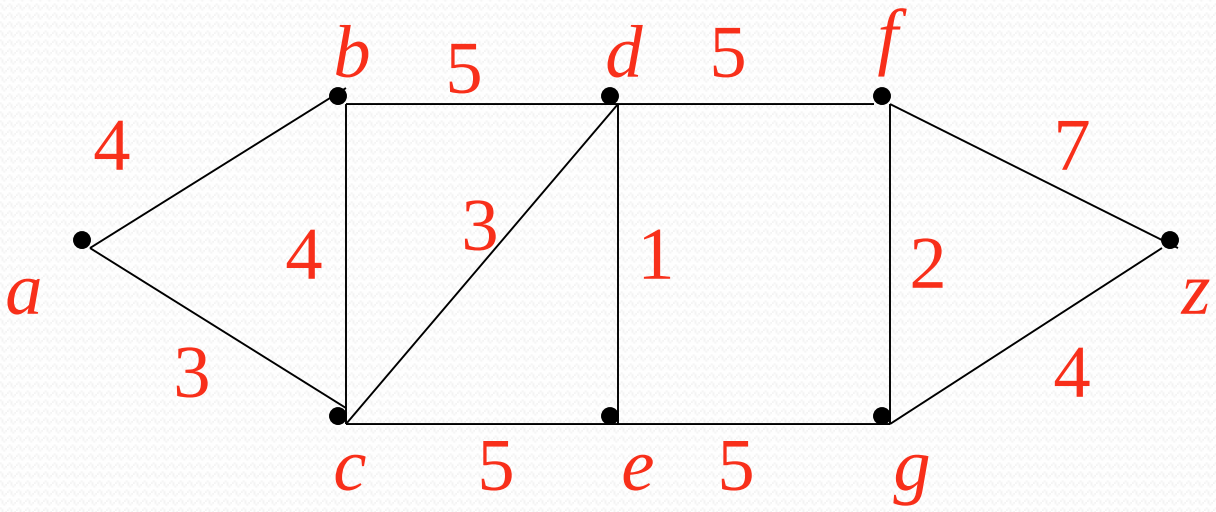


FIGURE 3 **A Weighted Simple Graph.**

What is the length of a shortest path between a and z in the weighted graph shown in Figure 3?

Problem: shortest path from a to z



The Traveling Salesman Problem

- The **traveling salesman problem** is one of the classical problems in computer science.
- A traveling salesman wants to visit a number of cities and then return to his starting point. Of course he wants to save time and energy, so he wants to determine the **shortest cycle** for his trip.
- We can represent the cities and the distances between them by a weighted, complete, undirected graph.
- The problem then is to find the **shortest cycle (of minimum total weight that visits each vertex exactly one)**.
- Finding the shortest cycle is different than Dijkstra's shortest path.
  It is much harder too, no polynomial time algorithm exists!

The Traveling Salesman Problem

- Importance:
  - Variety of scheduling application can be solved as a traveling salesmen problem.
  - Examples:
    - Ordering drill position on a drill press.
    - School bus routing.
  - The problem has theoretical importance because it represents a class of difficult problems known as NP-hard problems.

Travelling Salesman problem

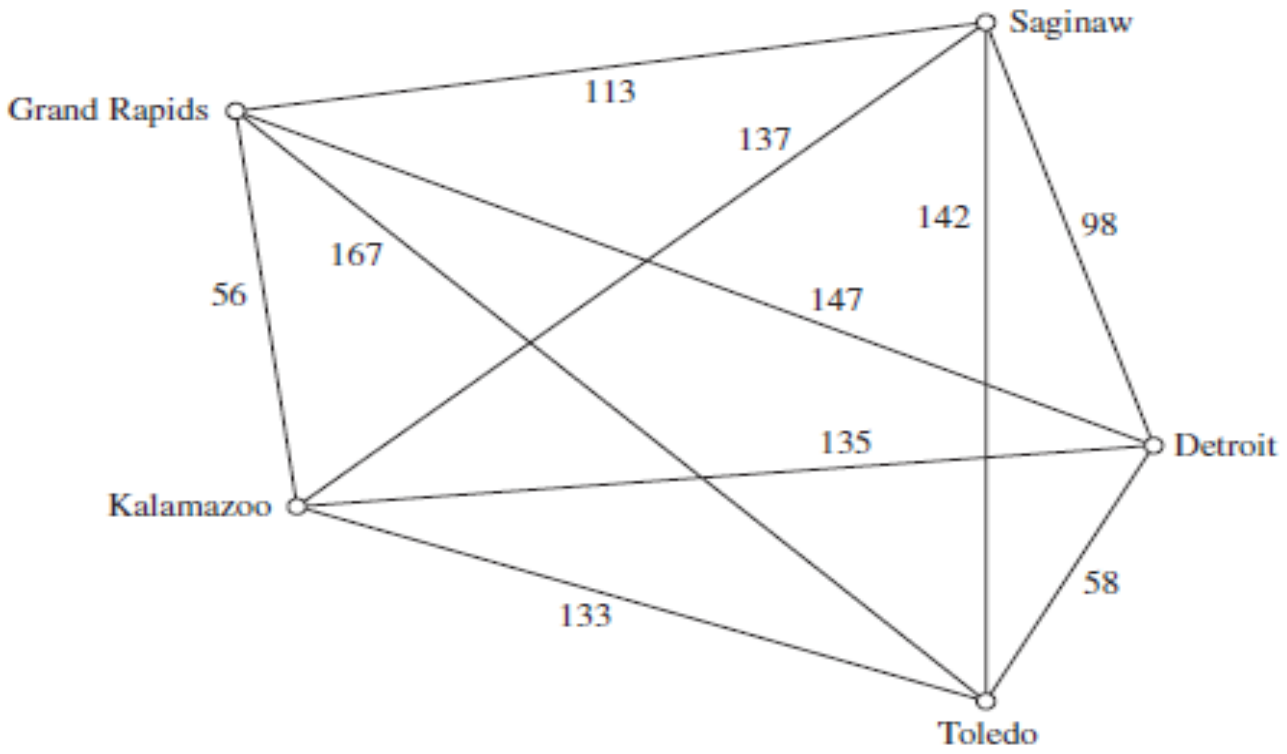


FIGURE 5 **The Graph Showing the Distances between Five Cities.**

Travelling Salesman problem

Route	*Total Distance (miles)*
Detroit–Toledo–Grand Rapids–Saginaw–Kalamazoo–Detroit	610
Detroit–Toledo–Grand Rapids–Kalamazoo–Saginaw–Detroit	516
Detroit–Toledo–Kalamazoo–Saginaw–Grand Rapids–Detroit	588
Detroit–Toledo–Kalamazoo–Grand Rapids–Saginaw–Detroit	458
Detroit–Toledo–Saginaw–Kalamazoo–Grand Rapids–Detroit	540
Detroit–Toledo–Saginaw–Grand Rapids–Kalamazoo–Detroit	504
Detroit–Saginaw–Toledo–Grand Rapids–Kalamazoo–Detroit	598
Detroit–Saginaw–Toledo–Kalamazoo–Grand Rapids–Detroit	576
Detroit–Saginaw–Kalamazoo–Toledo–Grand Rapids–Detroit	682
Detroit–Saginaw–Grand Rapids–Toledo–Kalamazoo–Detroit	646
Detroit–Grand Rapids–Saginaw–Toledo–Kalamazoo–Detroit	670
Detroit–Grand Rapids–Toledo–Saginaw–Kalamazoo–Detroit	728

Trees

Chapter 11

Chapter Summary

- Introduction to Trees
- Applications of Trees
- Tree Traversal
- Spanning Trees
- Minimum Spanning

Introduction to Trees

Section 11.1

Section Summary

- Introduction to Trees
- Rooted Trees
- Trees as Models
- Properties of Trees

Trees

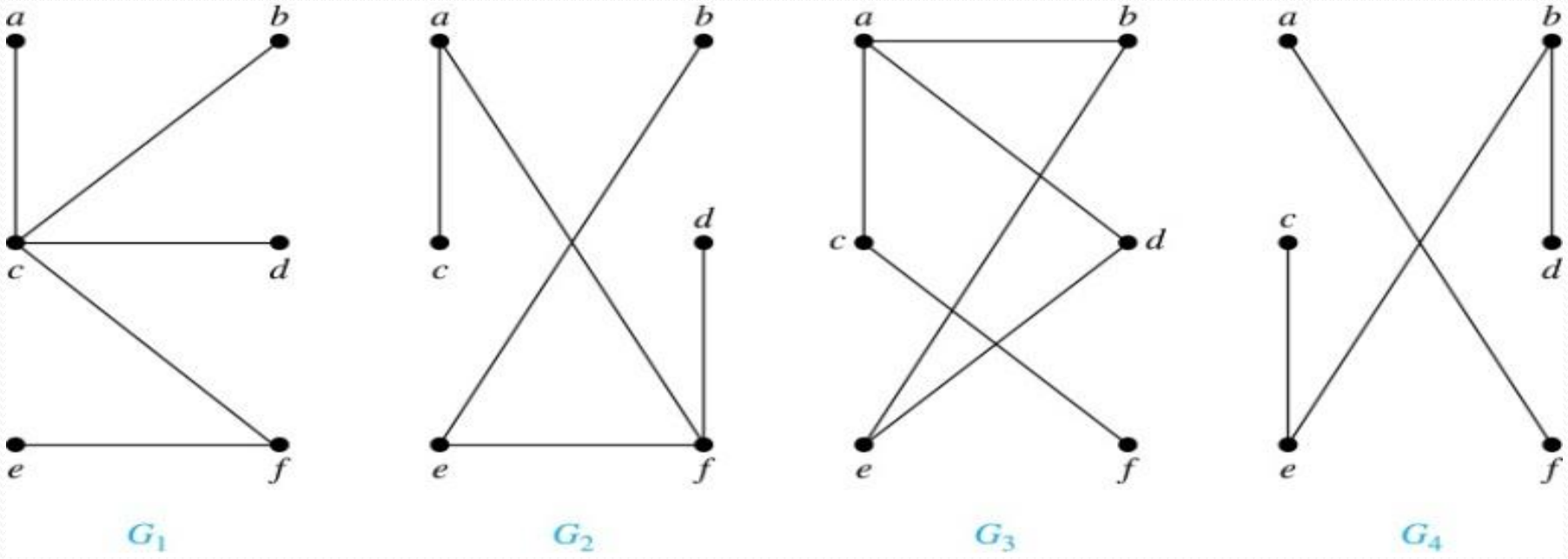
Definition: A *tree* is a connected undirected graph with no simple circuits.

Definition: An undirected graph is a tree if and only if there is a unique simple path between any two of its vertices. A tree cannot contain multiple edges or loops.

Definition: An undirected graph is a tree if and only if there is a unique simple path between any two of its vertices.

Trees

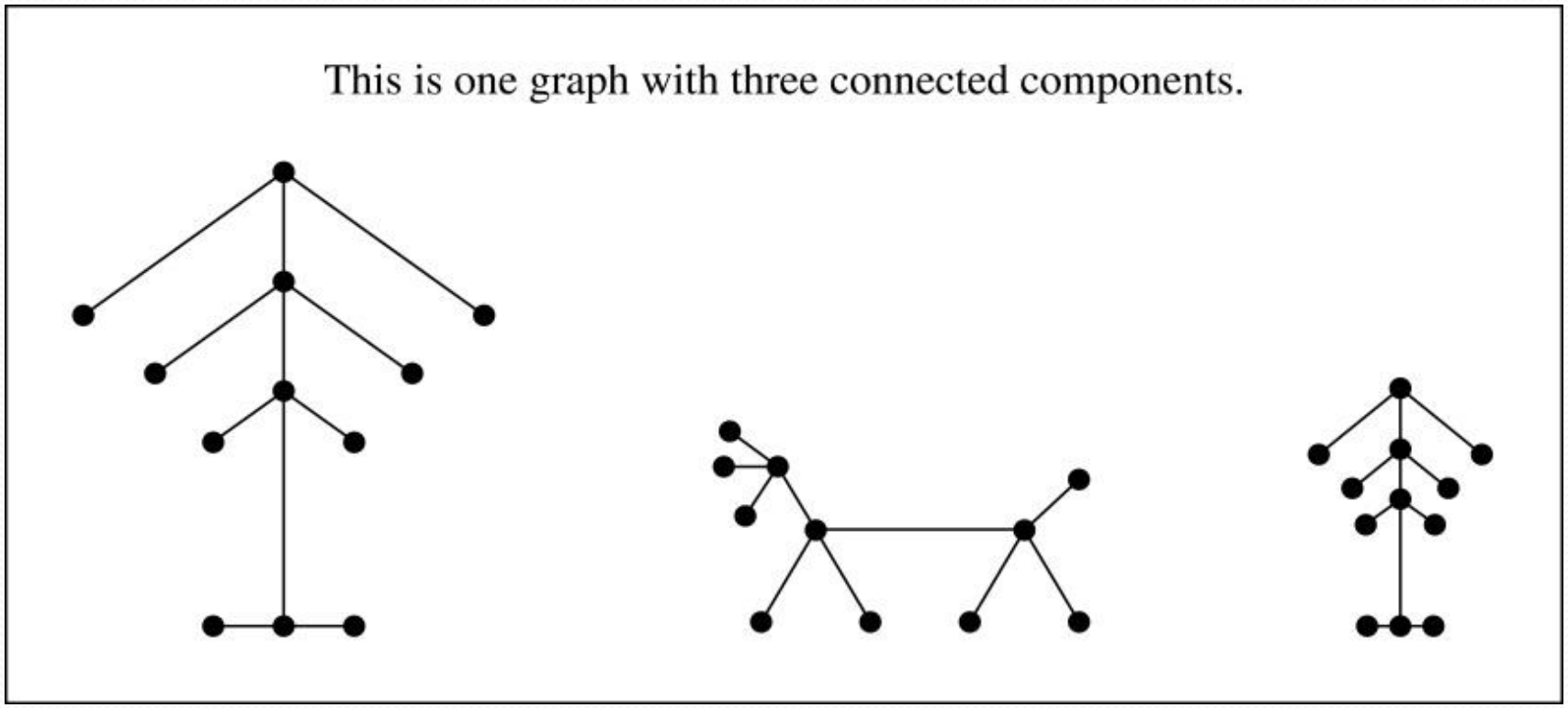
Example: Which of these graphs are trees?



Solution: G_1 and G_2 are trees - both are connected and have no simple circuits. G_3 is not a tree because e, b, a, d, e is a simple circuit,. G_4 is not a tree because it is not connected.

FOREST

Definition: A *forest* is a graph that has no simple circuit, but is not connected. Each of the connected components in a forest is a tree.

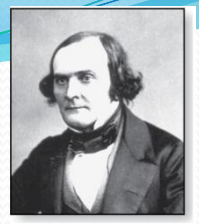


Applications of Trees

Section 11.2

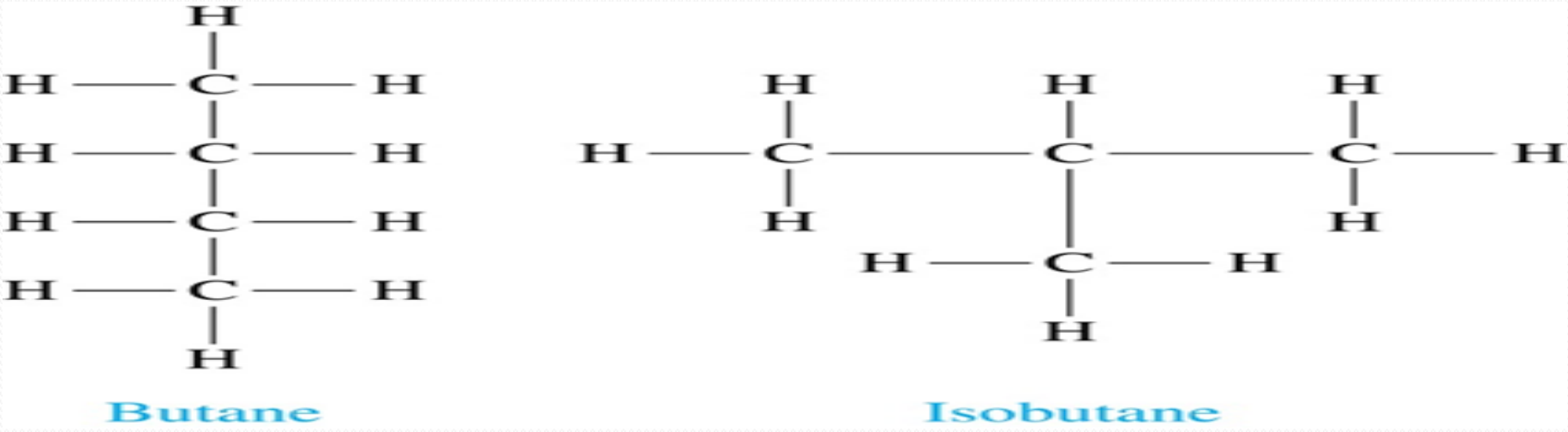
Trees as Models

Arthur Cayley
(1821-1895)



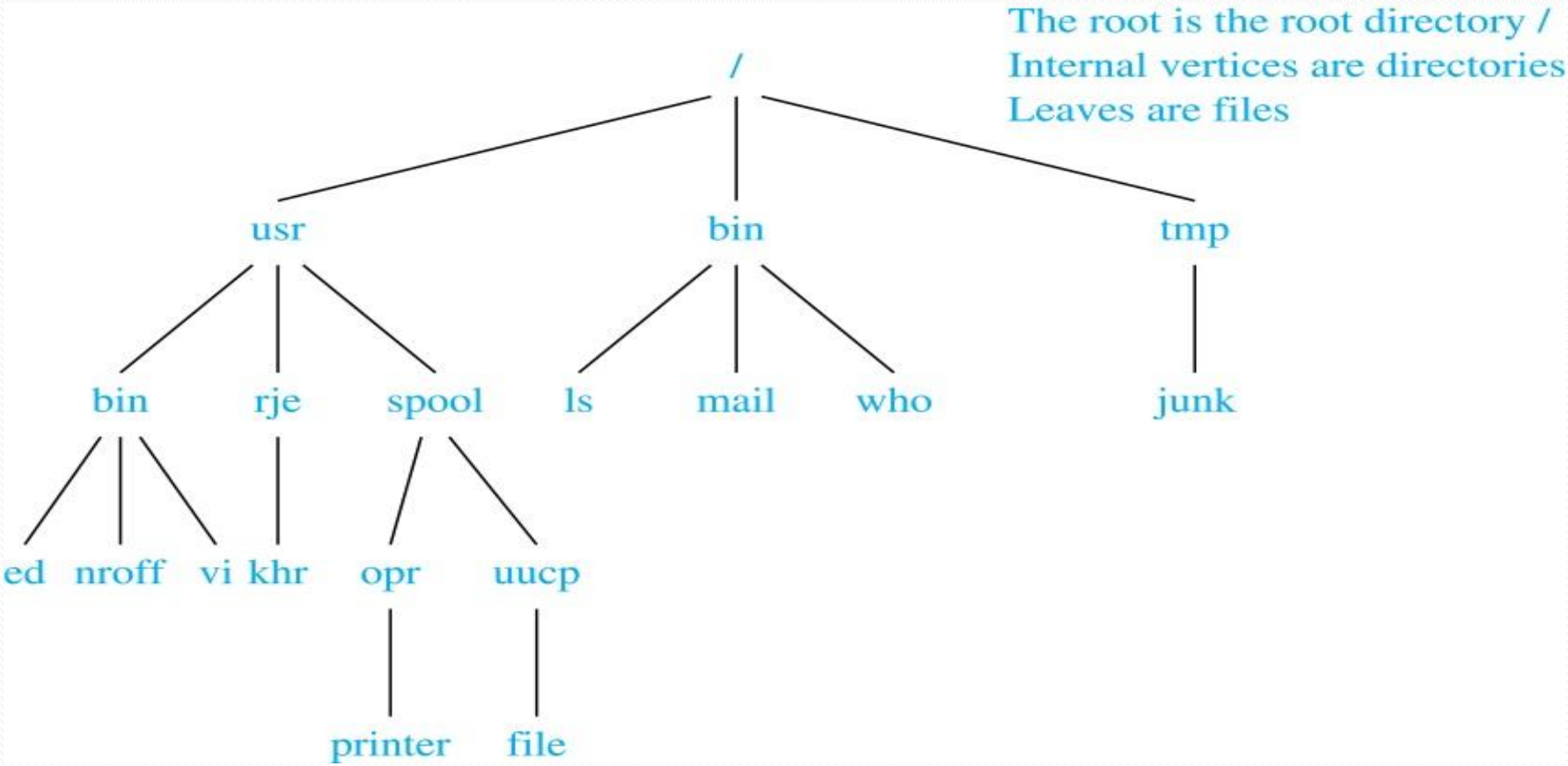
- Trees are used as models in computer science, chemistry, geology, botany, psychology, and many other areas.

- Trees were introduced by the mathematician Cayley in 1857 in his work counting the number of isomers of saturated hydrocarbons. The two isomers of butane are:



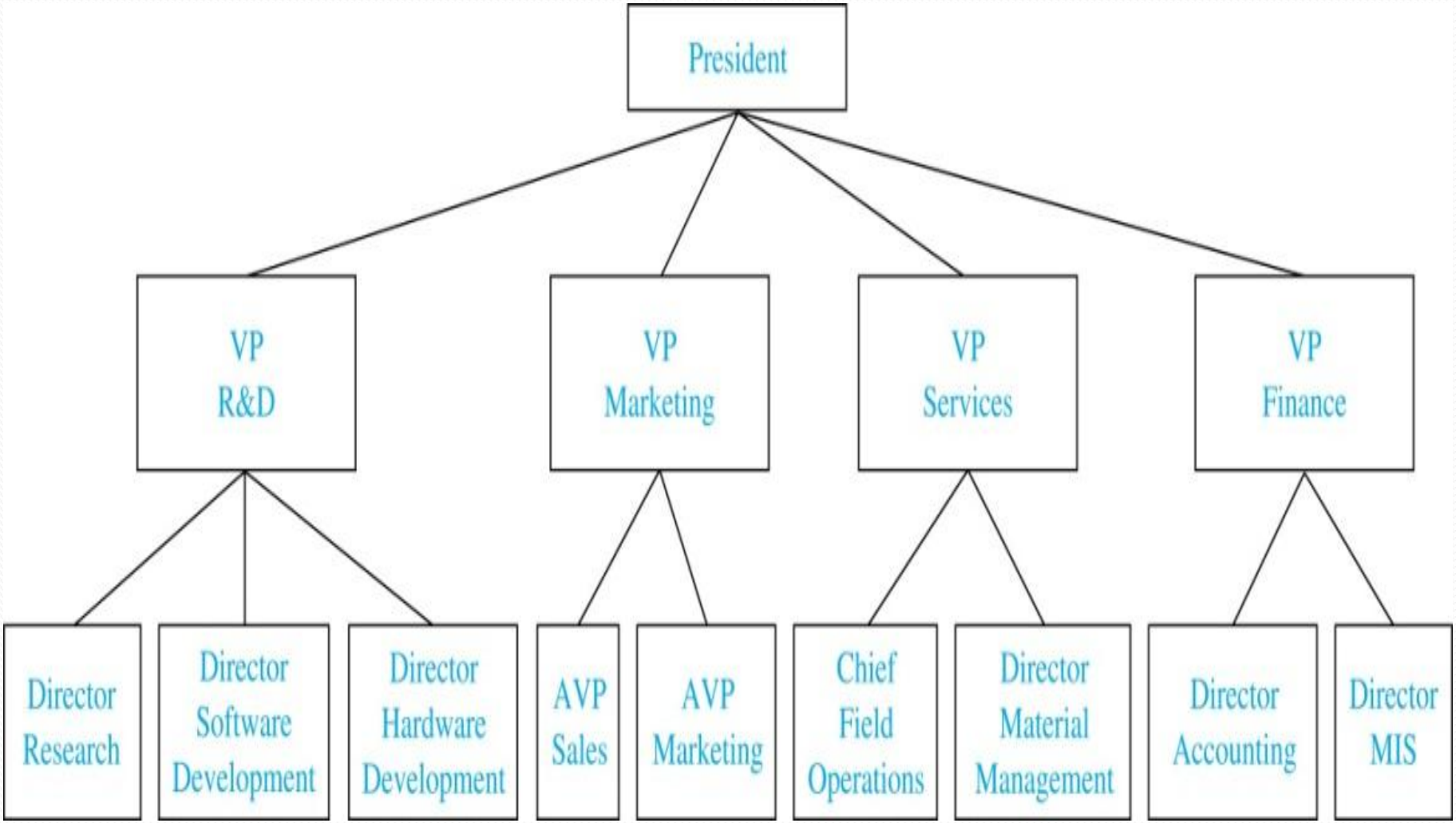
Trees as Models

- The organization of a computer file system into directories, subdirectories, and files is naturally represented as a tree.



Trees as Models

- Trees are used to represent the structure of organizations.

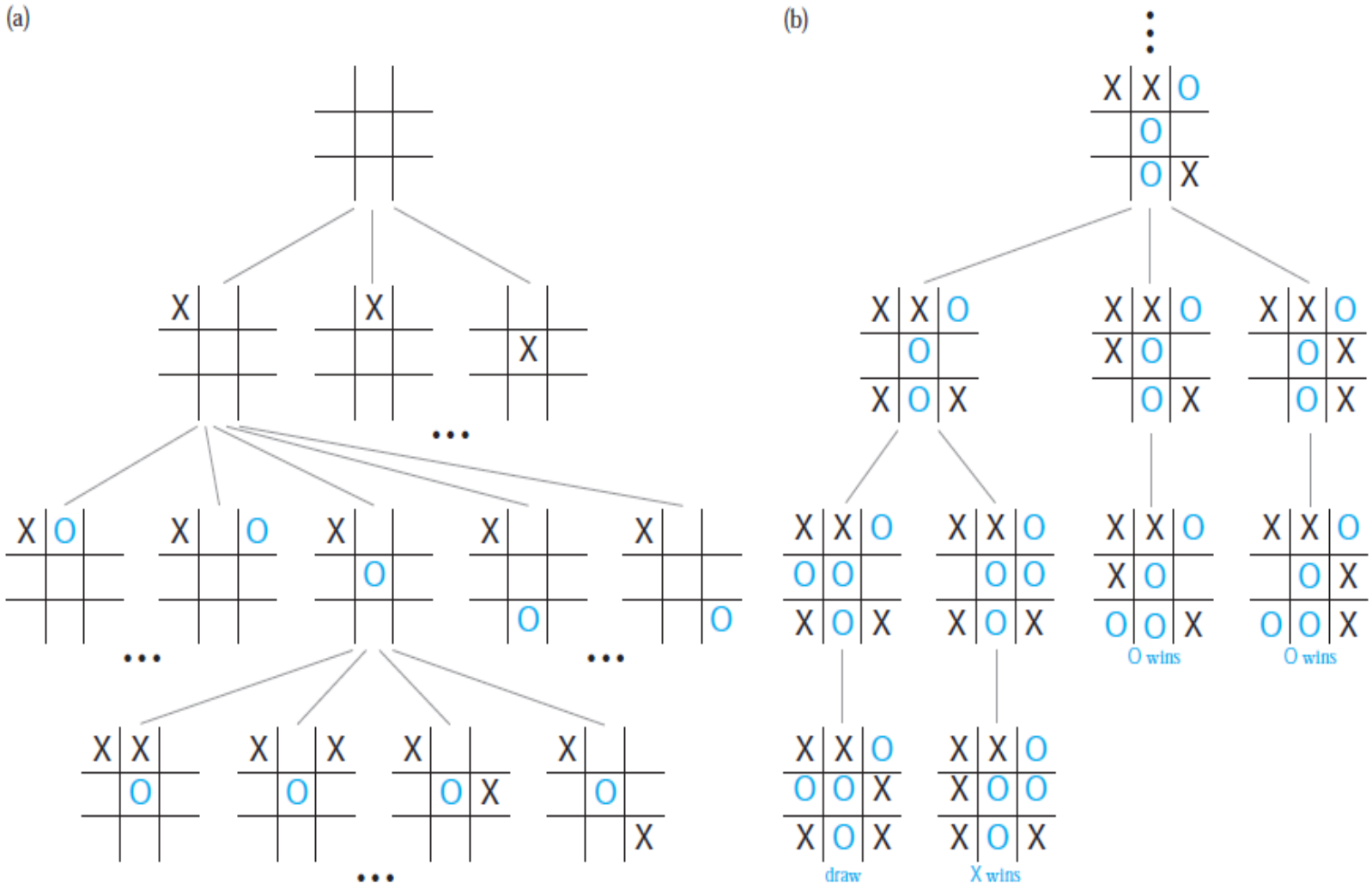


Applications of Trees

- **Game Trees**

Trees can be used to analyze certain types of games such as tic-tac-toe, nim, checkers, and chess.

Game Tree for Tic-Tac-Toe



Universal Address Systems

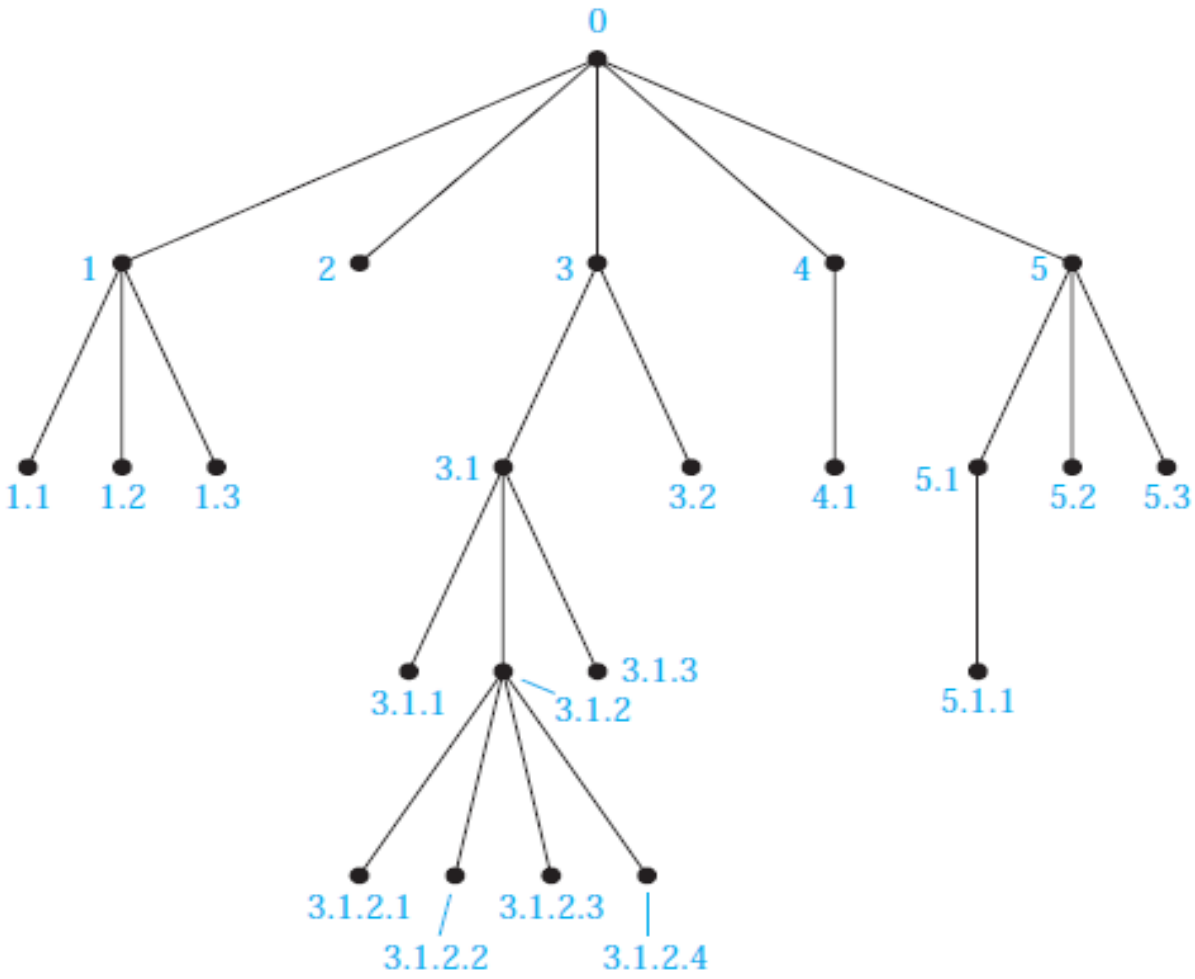


FIGURE 1 The Universal Address System of an Ordered Rooted Tree.

Prefix code

Definition: A code that has the property that the code of a character is never a prefix of the code of another character.

- A prefix code can be represented using a binary tree, where the characters are the labels of the leaves in the tree.
- The edges of the tree are labeled so that an edge leading to a left child is assigned a 0 and an edge leading to a right child is assigned a 1.
- The bit string used to encode a character is the sequence of labels of the edges in the unique path from the root to the leaf that has this character as its label.
- For instance, the tree in Figure 5 represents the encoding of *e* by 0, *a* by 10, *t* by 110, *n* by 1110, and *s* by 1111.

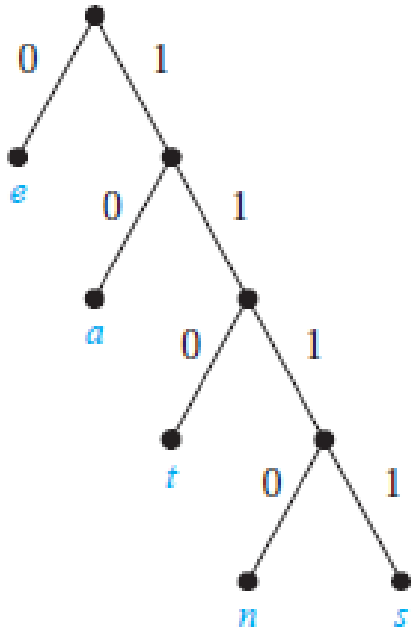


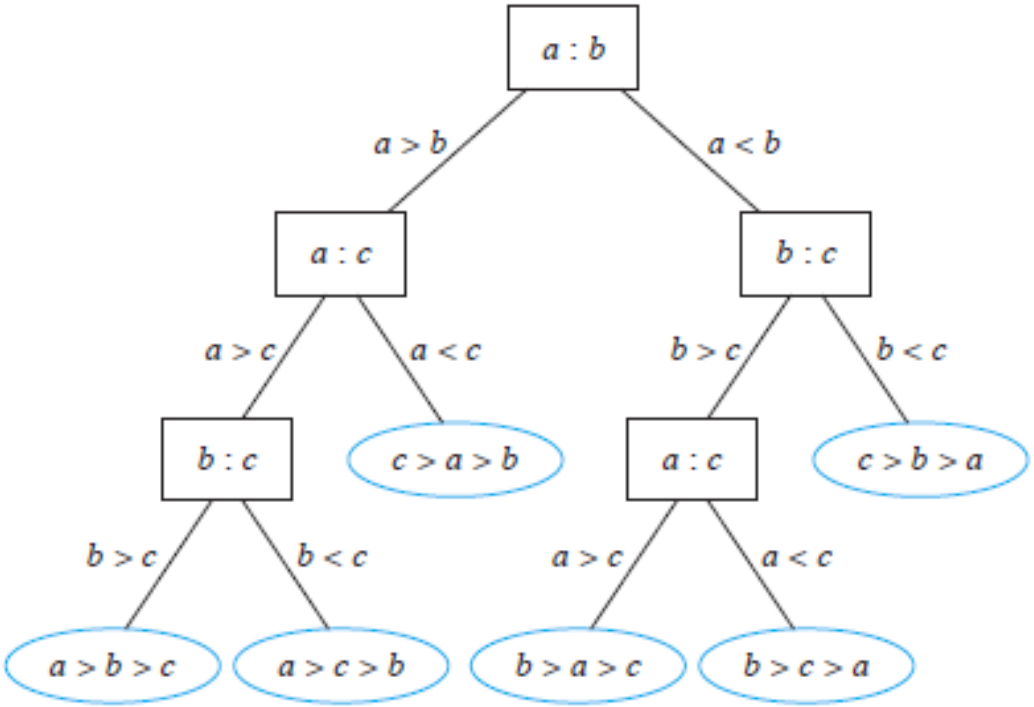
FIGURE 5 A Binary Tree with a Prefix Code.

Decision Trees

Definition: A rooted tree where each vertex represents a possible outcome of a decision and the leaves represent the possible solutions of a problem.

- Rooted trees can be used to model problems in which a series of decisions leads to a solution.
- The possible solutions of the problem correspond to the paths to the leaves of this rooted tree.

Example : A decision tree that orders the elements of the list *a*, *b*, *c*.

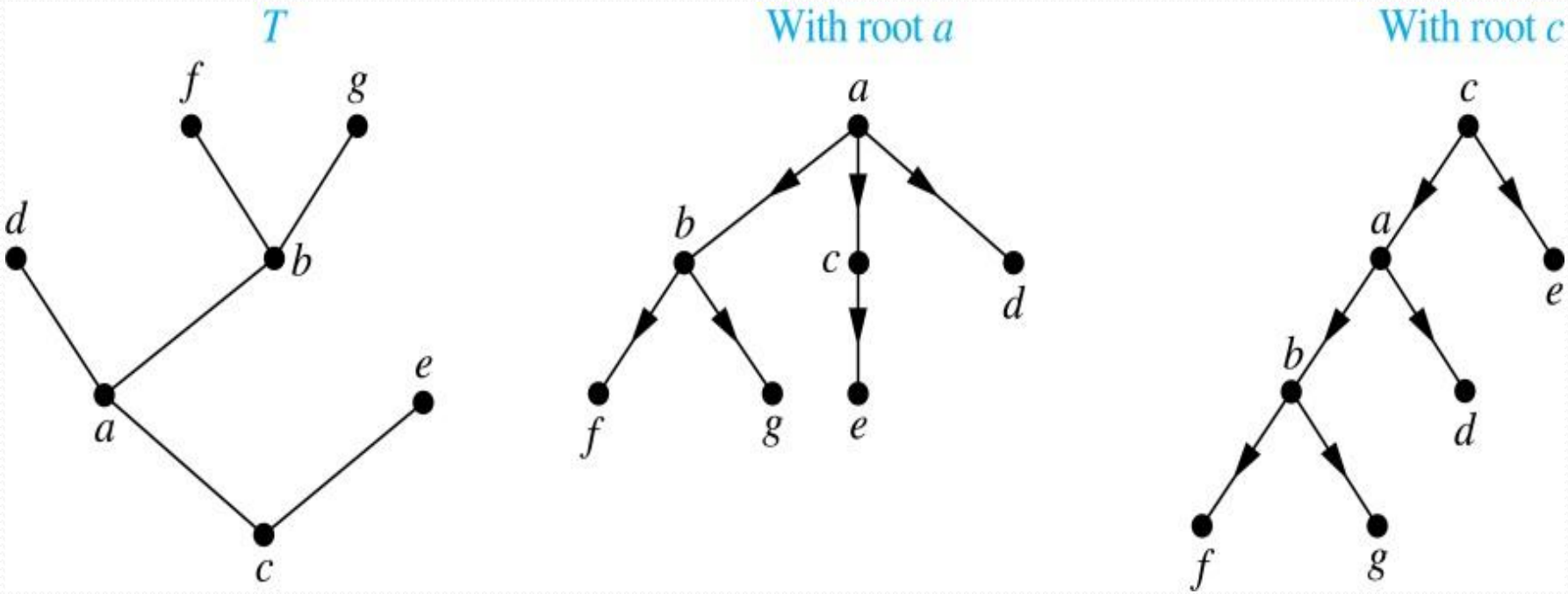


A Decision Tree for Sorting Three Distinct Elements.

Rooted Trees

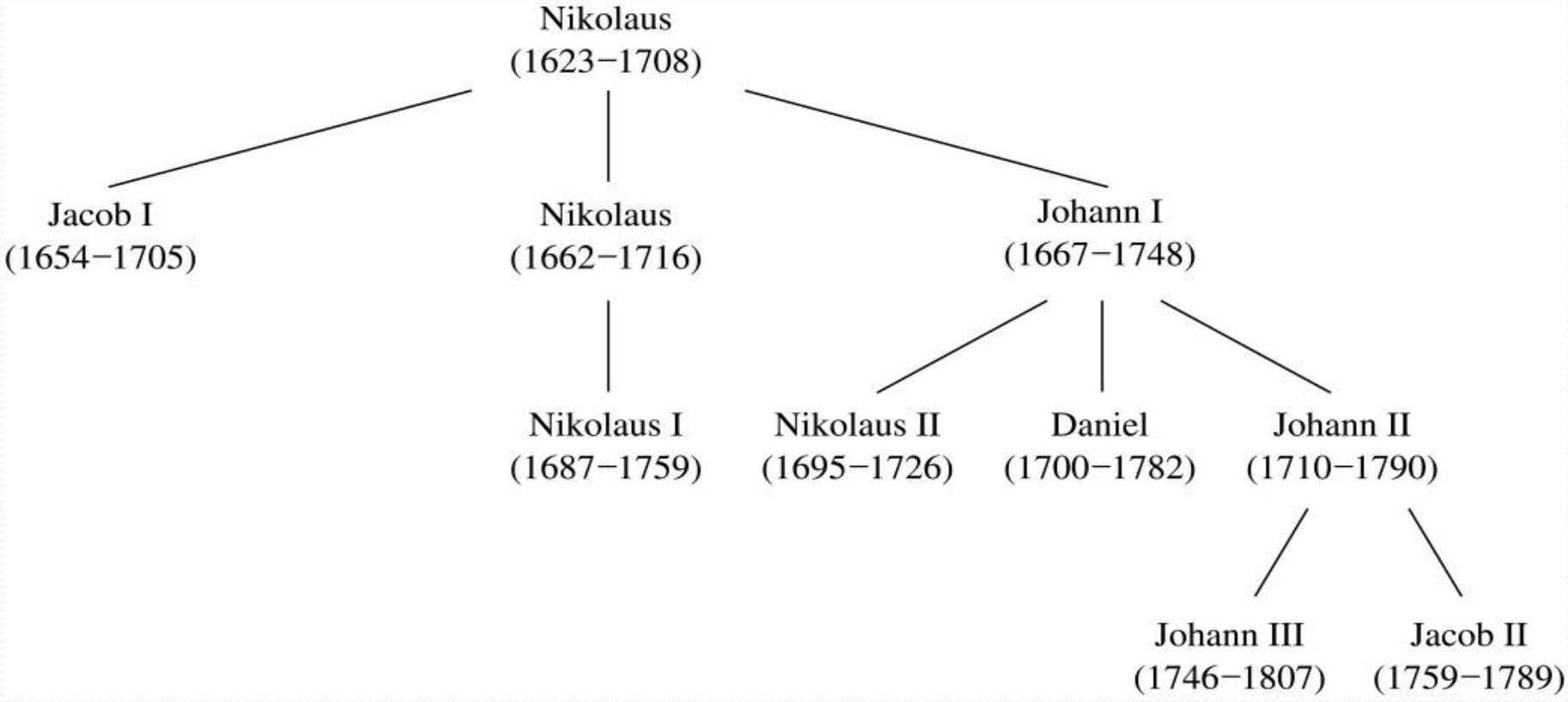
Definition: A *rooted tree* is a tree in which one vertex has been designated as the *root* and every edge is directed away from the root.

- An unrooted tree is converted into different rooted trees when different vertices are chosen as the root.



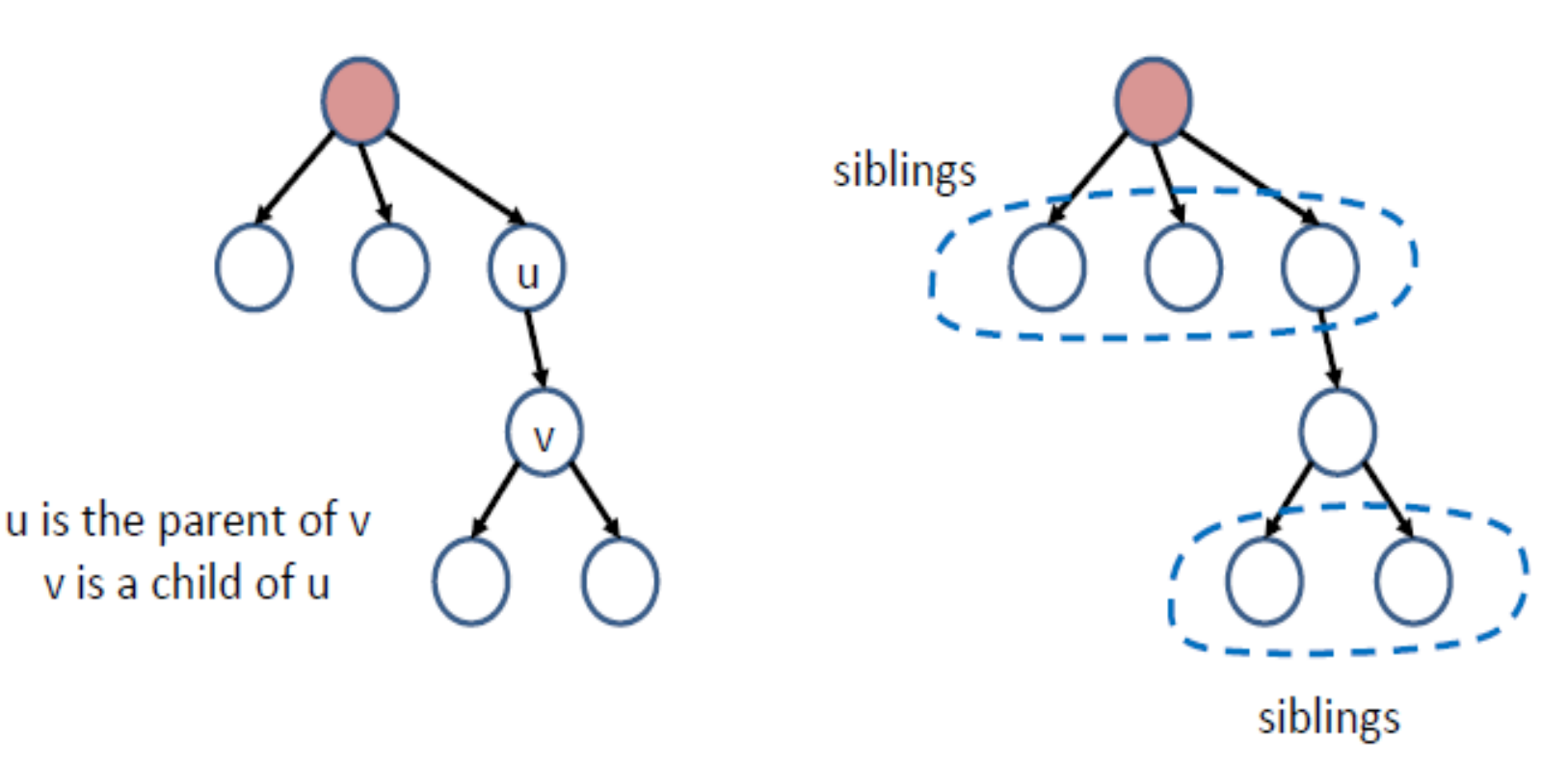
Rooted Tree Terminology

- Terminology for rooted trees is a mix from botany and genealogy (such as this family tree of the Bernoulli family of mathematicians).



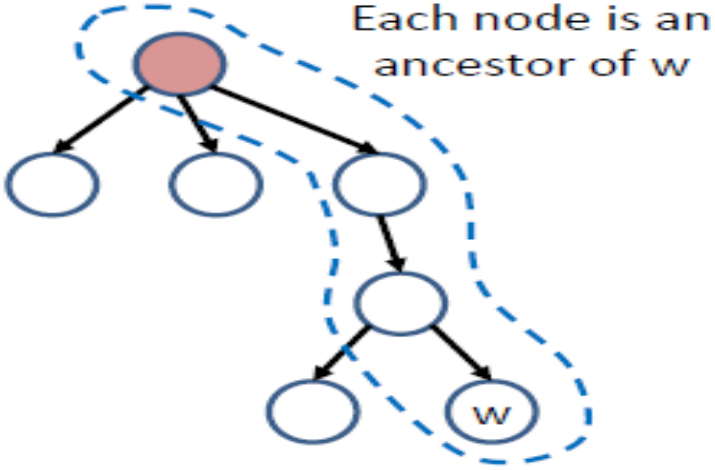
Rooted Tree Terminology

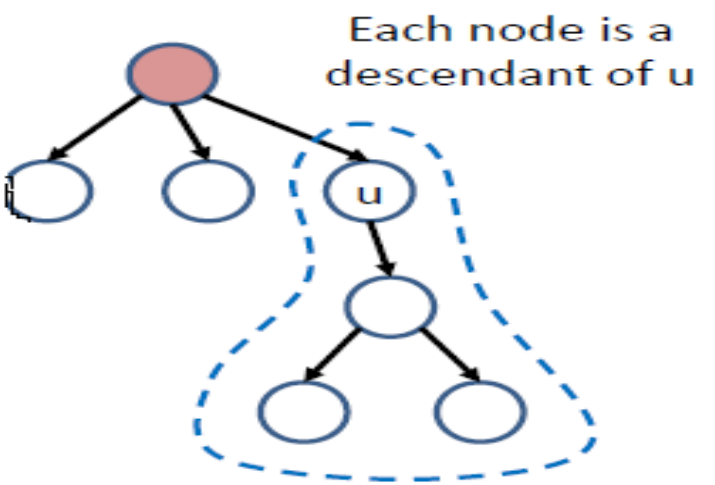
- If *v* is a vertex of a rooted tree other than the root, the *parent* of *v* is the unique vertex *u* such that there is a directed edge from *u* to *v*. When *u* is a parent of *v*, *v* is called a *child* of *u*. Vertices with the same parent are called *siblings*.



Rooted Tree Terminology

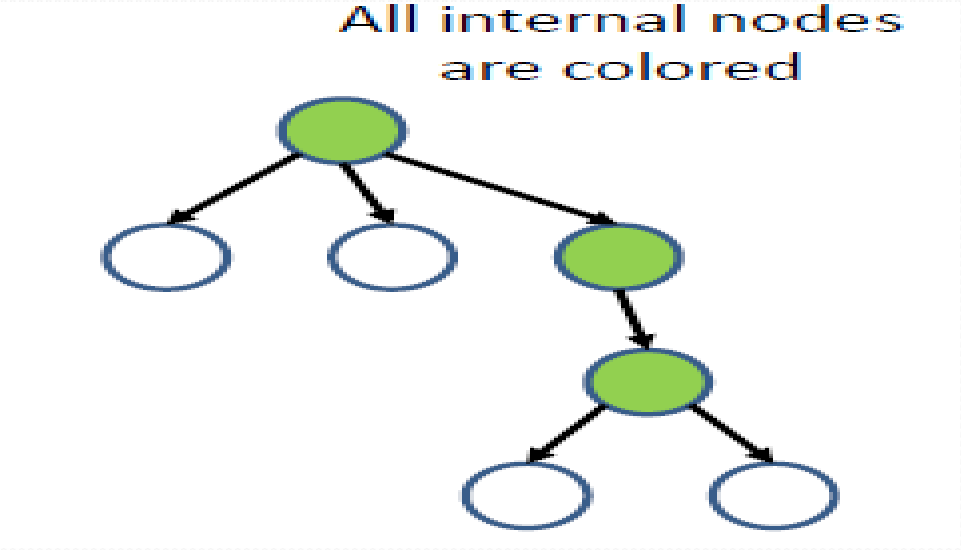
- The *ancestors* of a vertex are the vertices in the path from the root to this vertex, excluding the vertex itself and including the root.
- The *descendants* of a vertex *v* are those vertices that have *v* as an ancestor. The subtree rooted at u includes all the descendants of u, and all edges that connect between them.





Rooted Tree Terminology

- A vertex of a rooted tree with no children is called a *leaf*. Vertices that have children are called *internal vertices*.



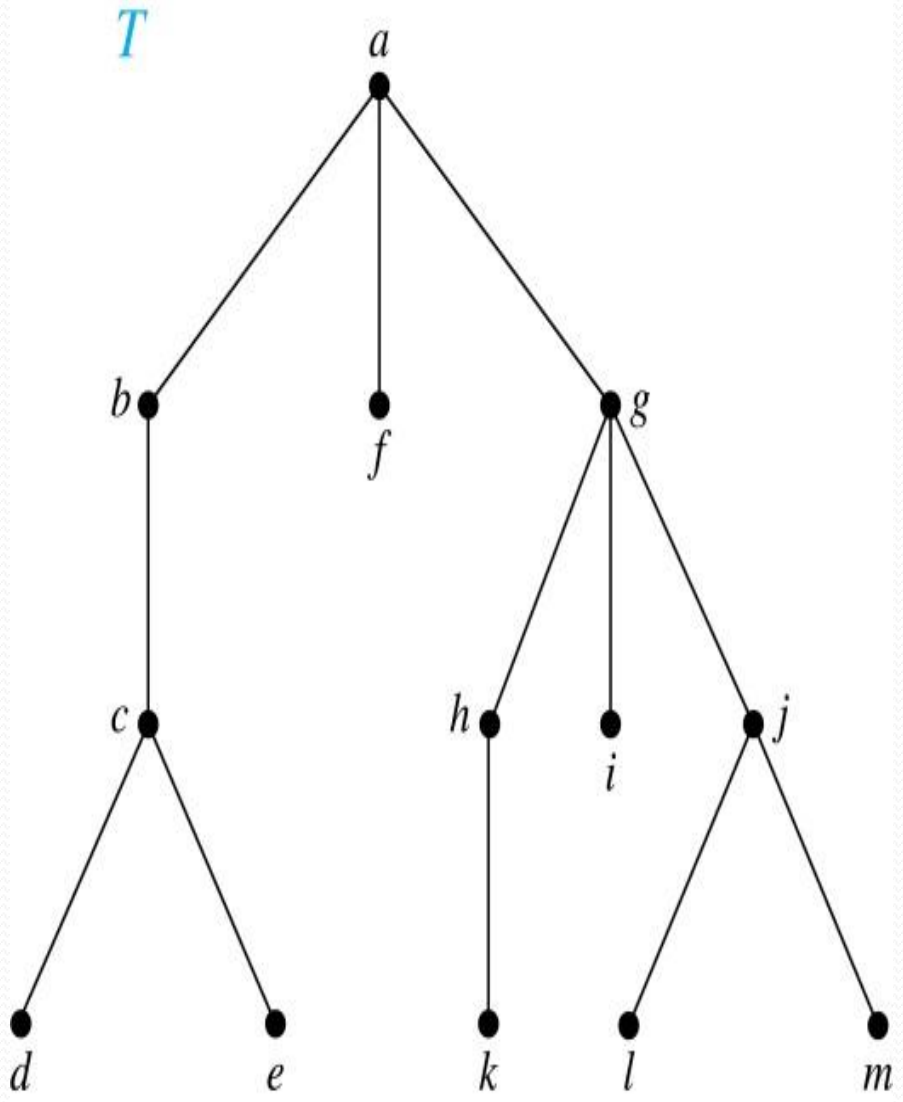
Terminology for Rooted Trees

Example: In the rooted tree T (with root a):

(i) Find the parent of c, the children of g, the siblings of h, the ancestors of e, and the descendants of b.

Solution:

(i) The parent of c is b. The children of g are h, i, and j. The siblings of h are i and j. The ancestors of e are c, b, and a. The descendants of b are c, d, and e.



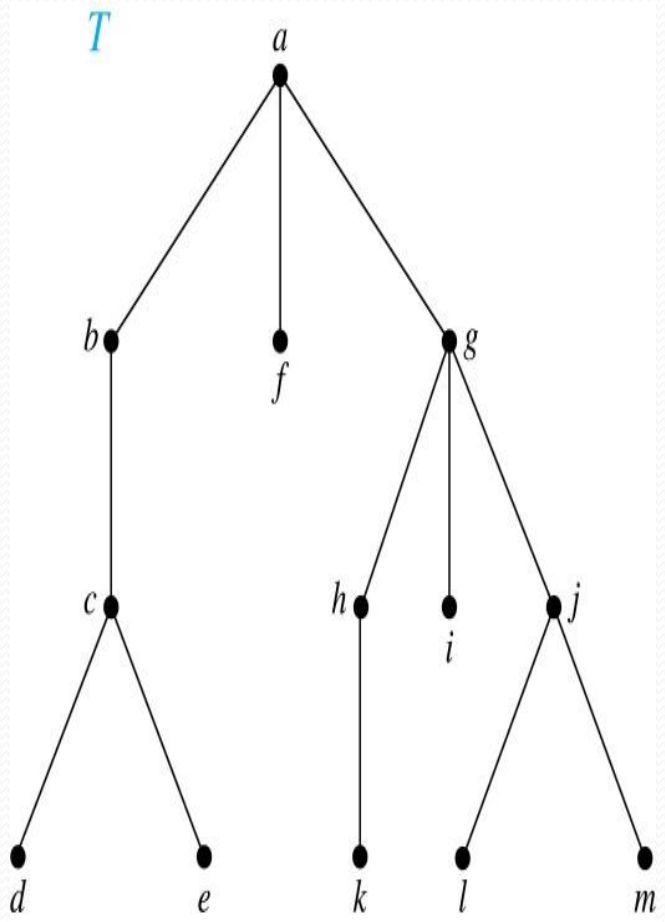
Terminology for Rooted Trees

Example: In the rooted tree T (with root a):

(i) Find all internal vertices and all leaves.

Solution:

(i) The internal vertices are a, b, c, g, h, and j. The leaves are d, e, f, i, k, l, and m.

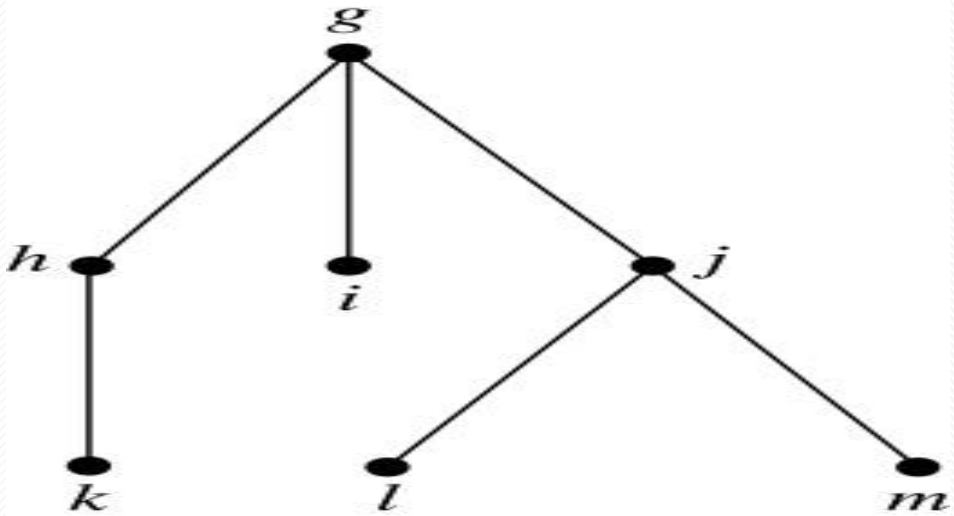


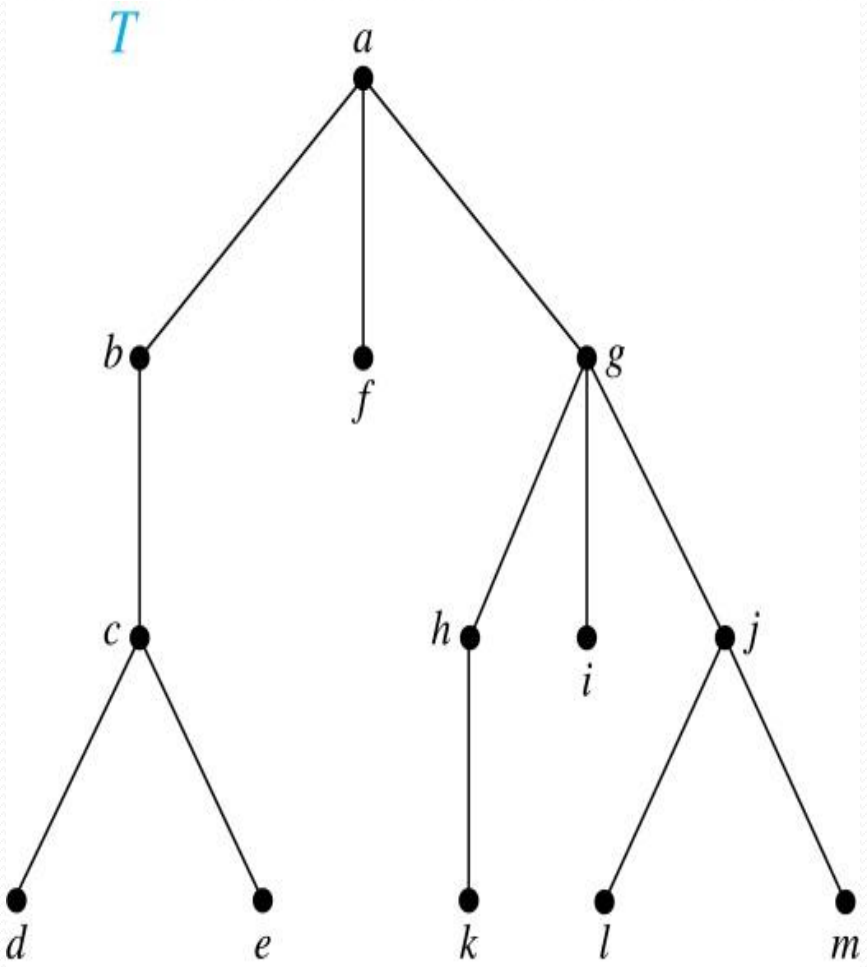
Terminology for Rooted Trees

(i) What is the subtree rooted at g?

Solution:

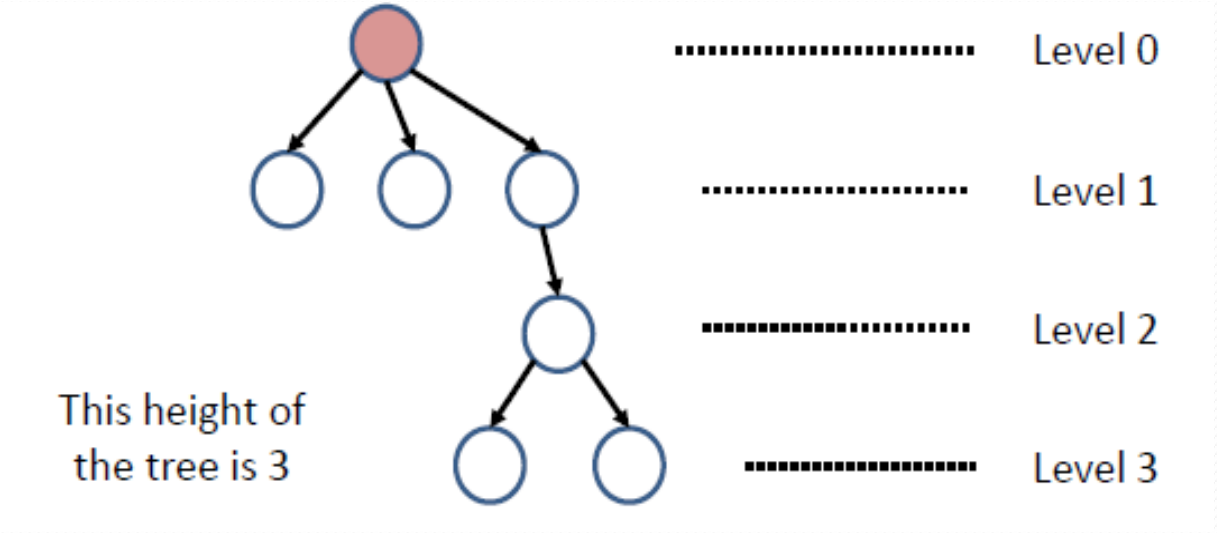
(i) We display the subtree rooted at g.





Level of vertices and height of trees

- When working with trees, we often want to have rooted trees where the sub trees at each vertex contain paths of approximately the same length.
- To make this idea precise we need some definitions:
  - The *level* of a vertex *v* in a rooted tree is the length of the unique path from the root to this vertex.
  - The *height* of a rooted tree is the maximum of the levels of the vertices.

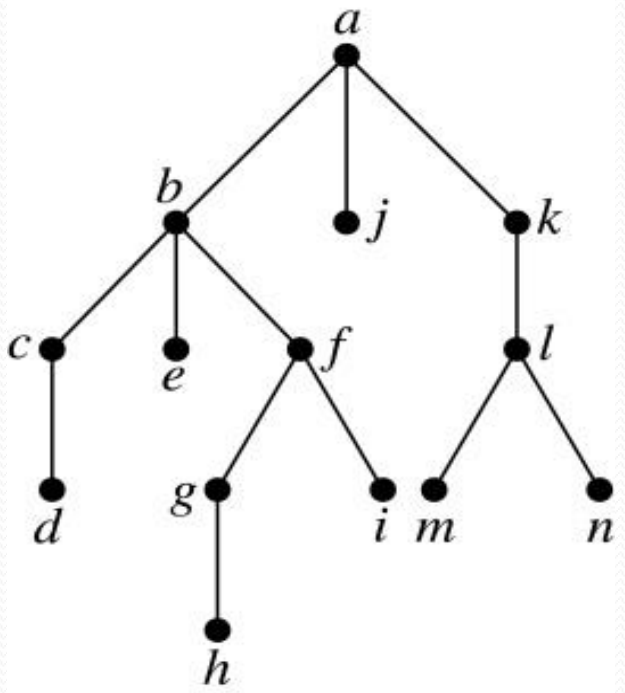


Level of vertices and height of trees

Example:

(*i*) Find the level of each vertex in the tree to the right.

(*ii*) What is the height of the tree?



Solution:

(*i*) The root *a* is at level 0.

Vertices *b*, *j*, and *k* are at level 1.

Vertices *c*, *e*, *f*, and *l* are at level 2.

Vertices *d*, *g*, *i*, *m*, and *n* are at level 3.

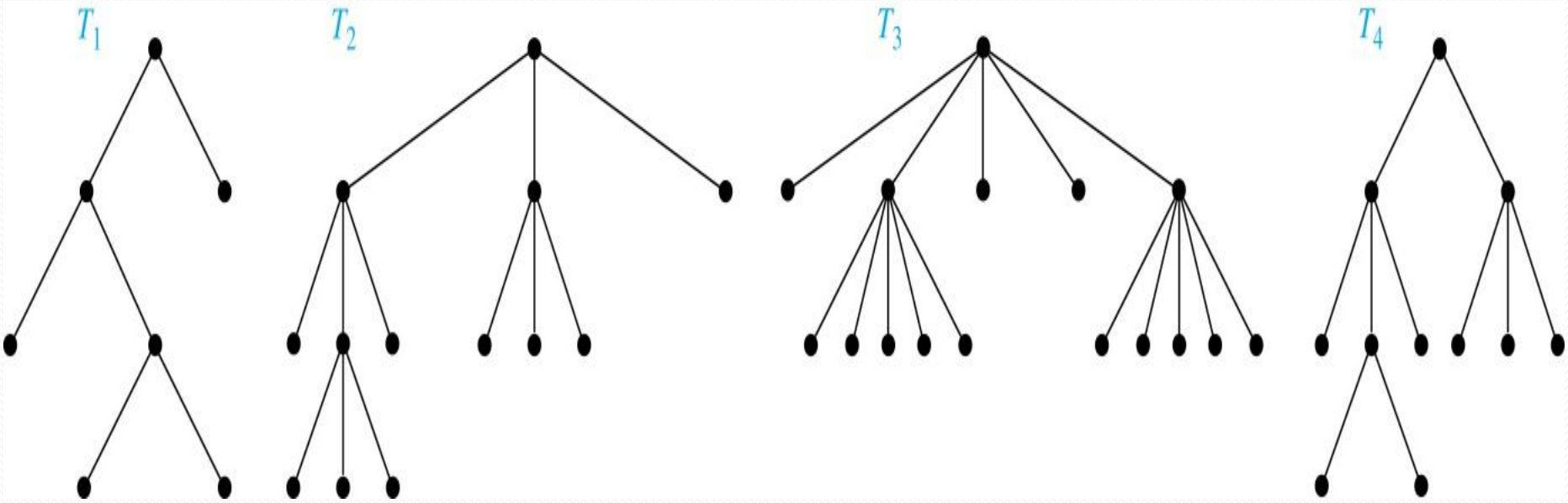
Vertex *h* is at level 4.

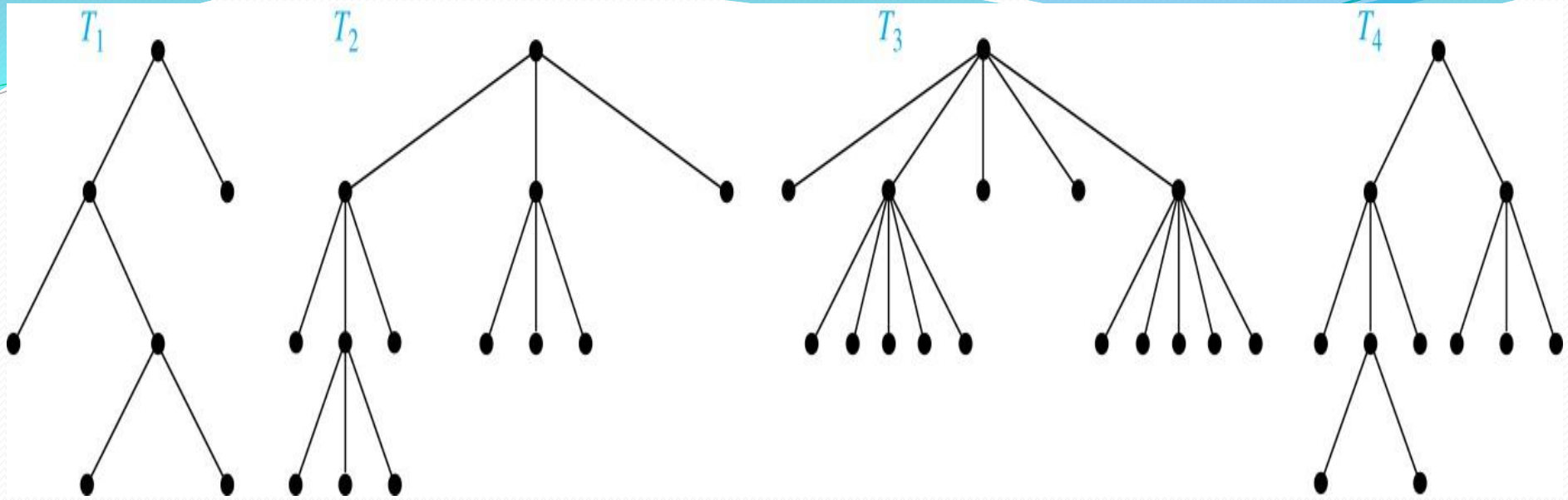
(*ii*) The height is 4, since 4 is the largest level of any vertex.

m-ary Rooted Trees

Definition: A rooted tree is called an *m-ary tree* if every internal vertex has no more than *m* children. The tree is called a *full m-ary tree* if every internal vertex has exactly *m* children. An *m*-ary tree with $m = 2$ is called a *binary* tree.

Example: Are the following rooted trees full *m*-ary trees for some positive integer *m*?





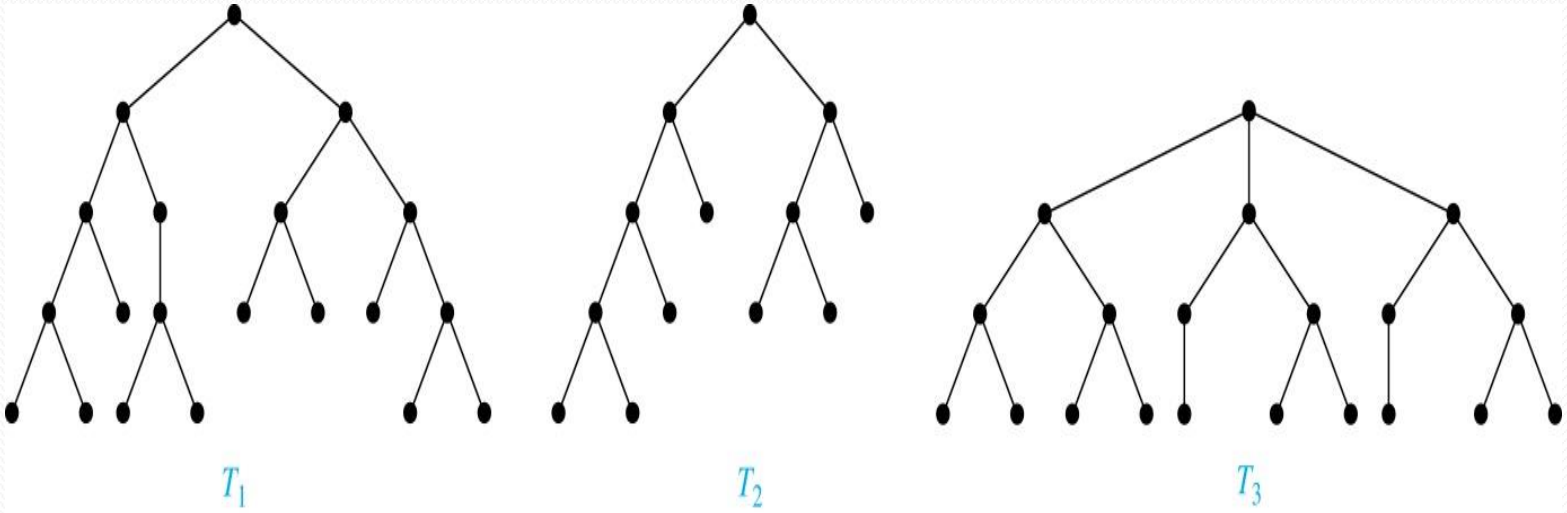
Solution:

- T_1 is a full binary tree because each of its internal vertices has two children.
- T_2 is a full 3-ary tree because each of its internal vertices has three children.
- In T_3 each internal vertex has five children, so T_3 is a full 5-ary tree.
- T_4 is not a full m-ary tree for any m because some of its internal vertices have two children and others have three children.

Balanced *m*-Ary Trees

Definition: A rooted *m*-ary tree of height *h* is *balanced* if all leaves are at levels *h* or $h - 1$.

Example: Which of the rooted trees shown below is balanced?

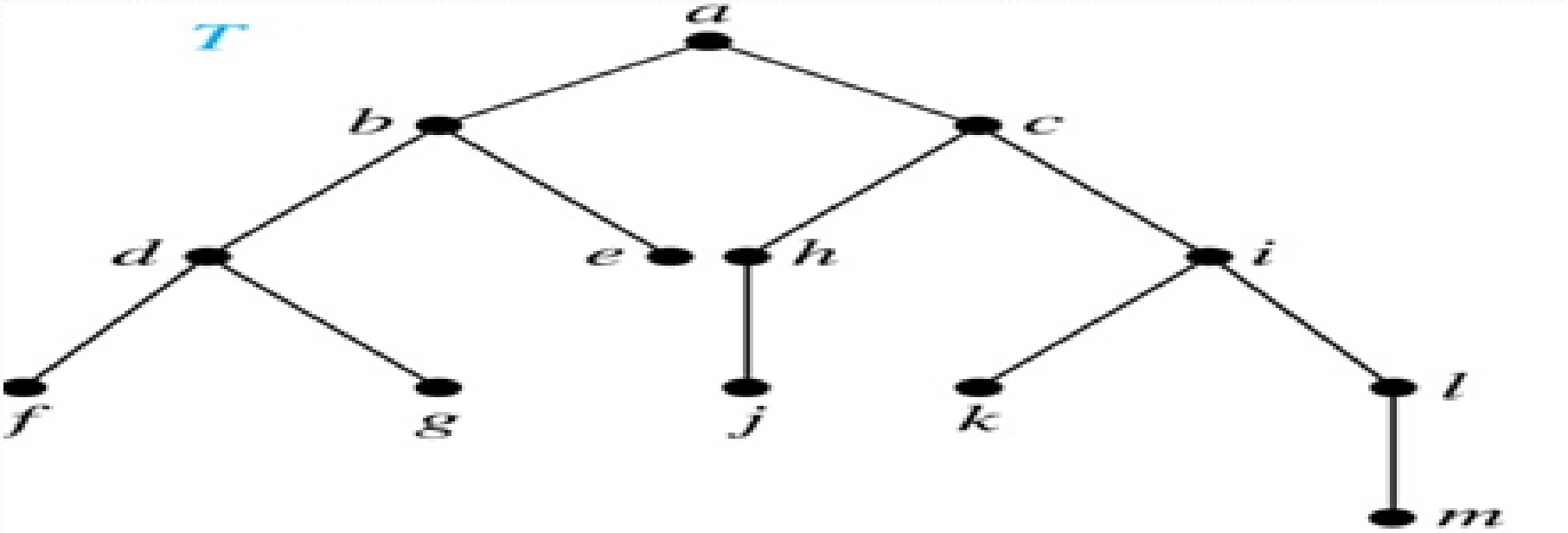


Solution: T_1 and T_3 are balanced, but T_2 is not because it has leaves at levels 2, 3, and 4.

Ordered Rooted Trees

Definition: An *ordered rooted tree* is a rooted tree where the children of each internal vertex are ordered.

- We draw ordered rooted trees so that the children of each internal vertex are shown in order from left to right.



Binary Trees

Definition: A *binary tree* is an ordered rooted where each internal vertex has at most two children. If an internal vertex of a binary tree has two children, the first is called the *left child* and the second the *right child*. The tree rooted at the left child of a vertex is called the *left subtree* of this vertex, and the tree rooted at the right child of a vertex is called the *right subtree* of this vertex.

Example:

Consider the binary tree T.

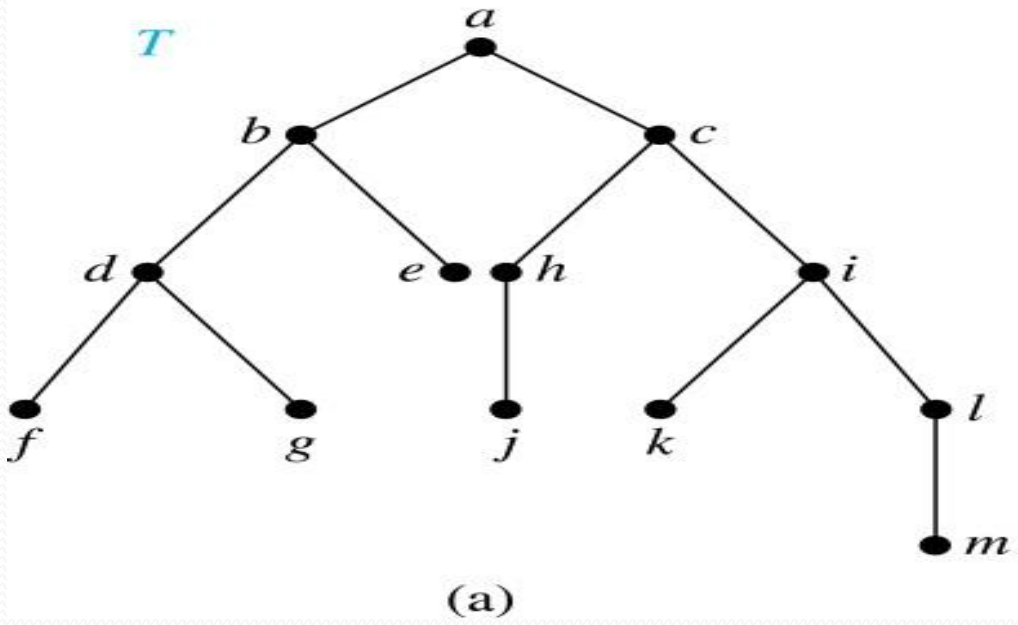
(*i*) What are the left and right children of d?

(*ii*) What are the left and right sub trees of c?

Solution:

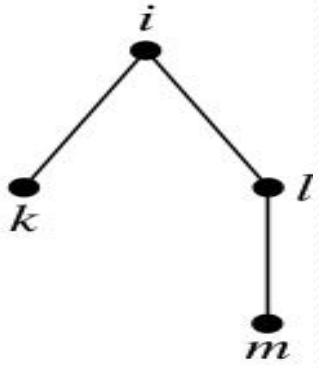
(*i*) The left child of d is f and the right child is g.

(*ii*) The left and right subtrees of c are displayed in (b) and (c).



(a)

(b)



(c)

Properties of Trees

- A tree with n vertices has $n - 1$ edges.
- A full m-ary tree with i internal vertices has $n = mi + 1$ vertices.
- A full m-ary tree with:

  (i) *n vertices has $i = (n - 1)/m$ internal vertices and $l = [(m - 1)n + 1]/m$ leaves,*

  (ii) i internal vertices has $n = mi + 1$ vertices and $l = (m - 1)i + 1$ leaves,

  (iii) l leaves has $n = (ml - 1)/(m - 1)$ vertices and $i = (l - 1)/(m - 1)$ internal vertices.

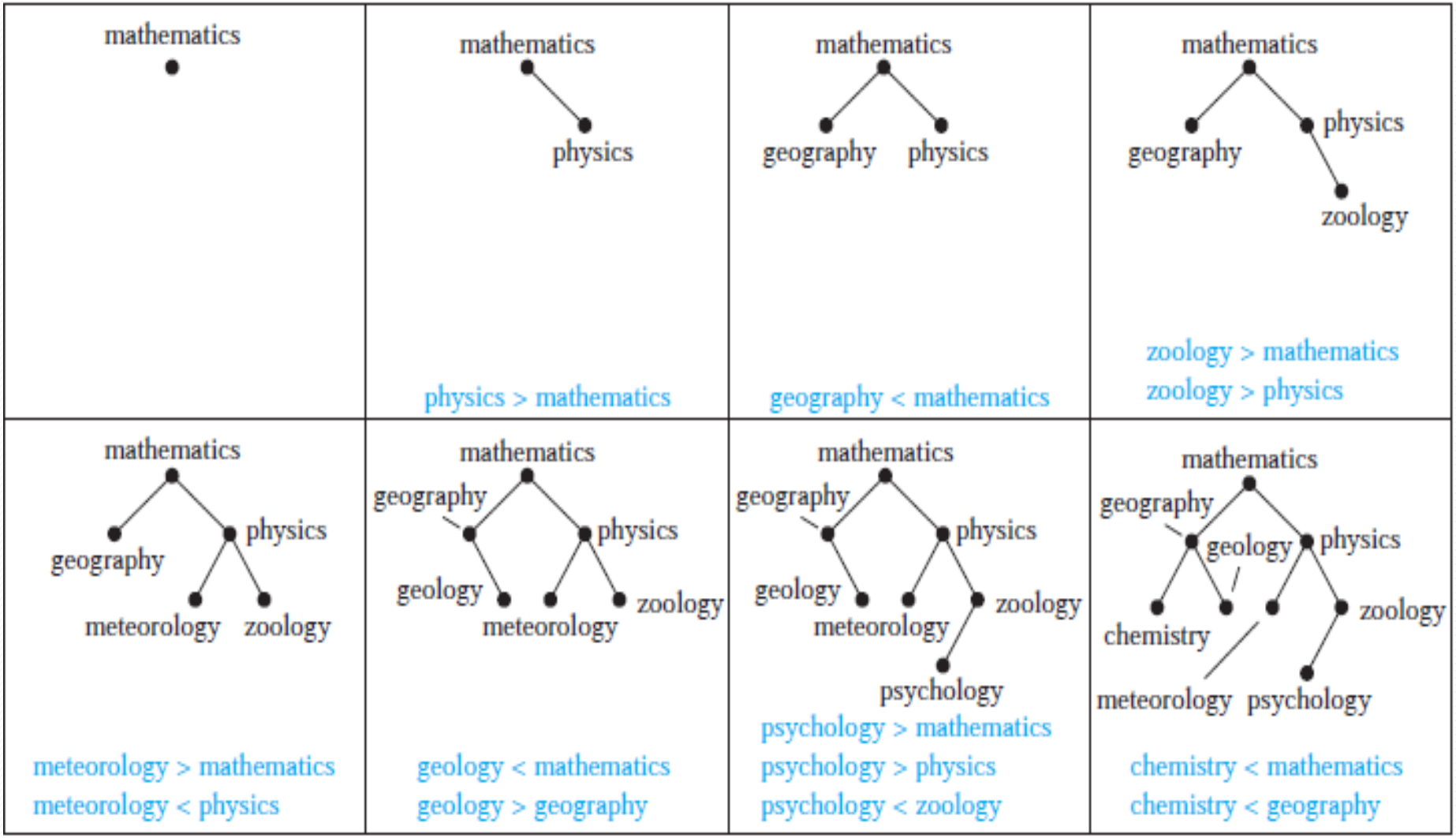
- There are at most m^h leaves in an m-ary tree of height h.

Binary Search Tree

Definition: A binary tree in which the vertices are labeled with items so that a label of a vertex is greater than the labels of all vertices in the left subtree of this vertex and is less than the labels of all vertices in the right subtree of this vertex.

- Searching for items in a list is one of the most important tasks that arises in computer science.
- Our primary goal is to implement a searching algorithm that finds items efficiently when the items are totally ordered. This can be accomplished through the use of a binary search tree.

Example : Form a binary search tree for the words mathematics, physics, geography, zoology, meteorology, geology, psychology, and chemistry (using alphabetical order).



Tree Traversal

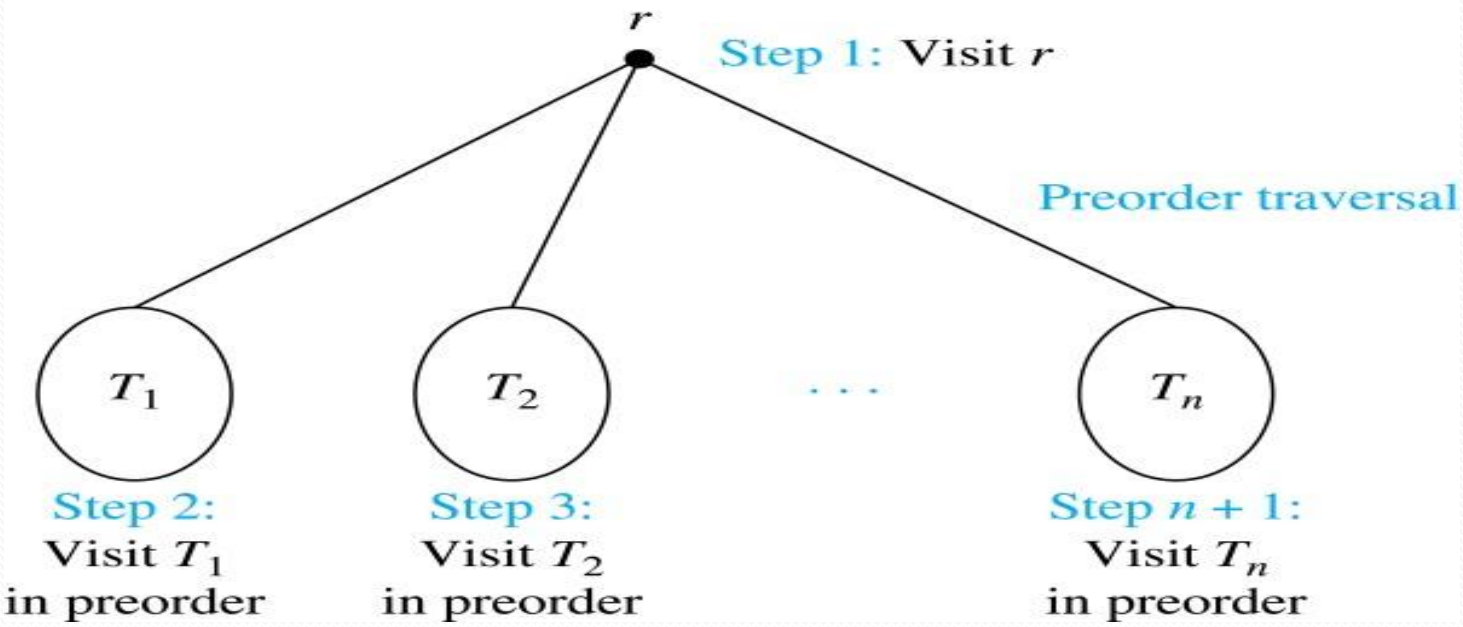
Section 11.3

Tree Traversal

- Procedures for systematically visiting every vertex of an ordered tree are called *traversals*.
- The three most commonly used *traversals* are *preorder traversal*, *inorder traversal*, and *postorder traversal*.

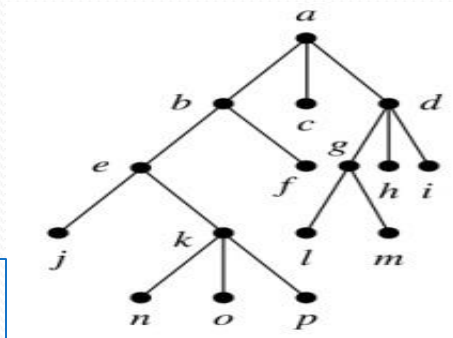
Preorder Traversal

Definition: Let T be an ordered rooted tree with root r. If T consists only of r, then r is the *preorder traversal* of T. Otherwise, suppose that T_1, T_2, ..., T_n are the subtrees of r from left to right in T. The preorder traversal begins by visiting r, and continues by traversing T_1 in preorder, then T_2 in preorder, and so on, until T_n is traversed in preorder.

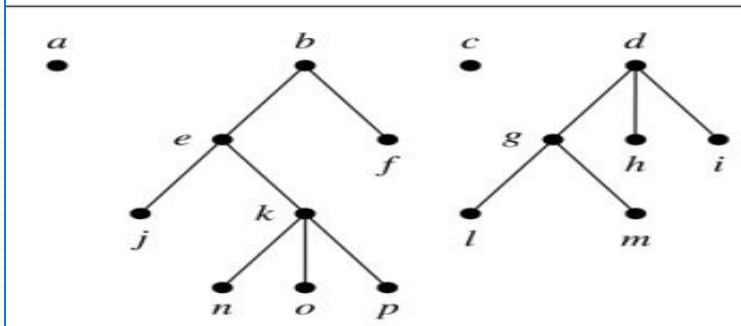


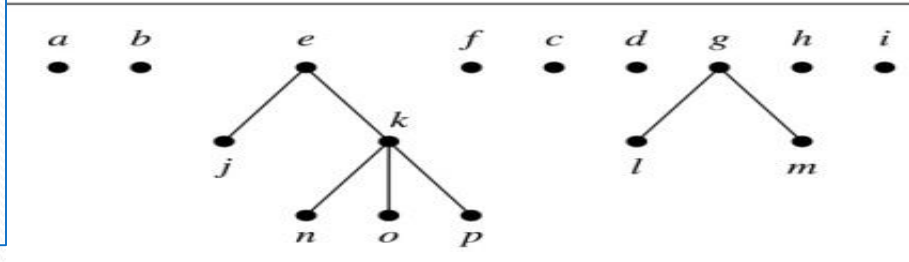
Preorder Traversal (*continued*)

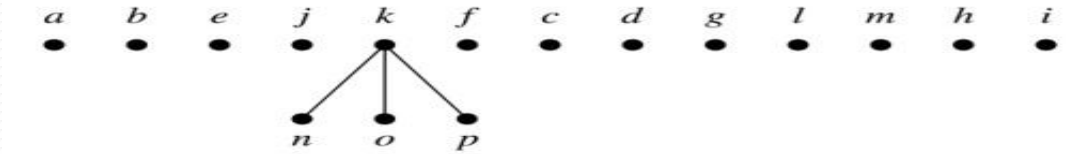
procedure *preorder* (*T*: ordered rooted tree)
r := root of *T*
list *r*
for each child *c* of *r* from left to right
    T(*c*) := subtree with *c* as root
    preorder(*T*(*c*))

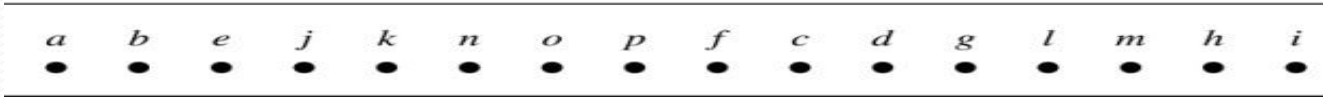


Preorder traversal: Visit root, visit subtrees left to right



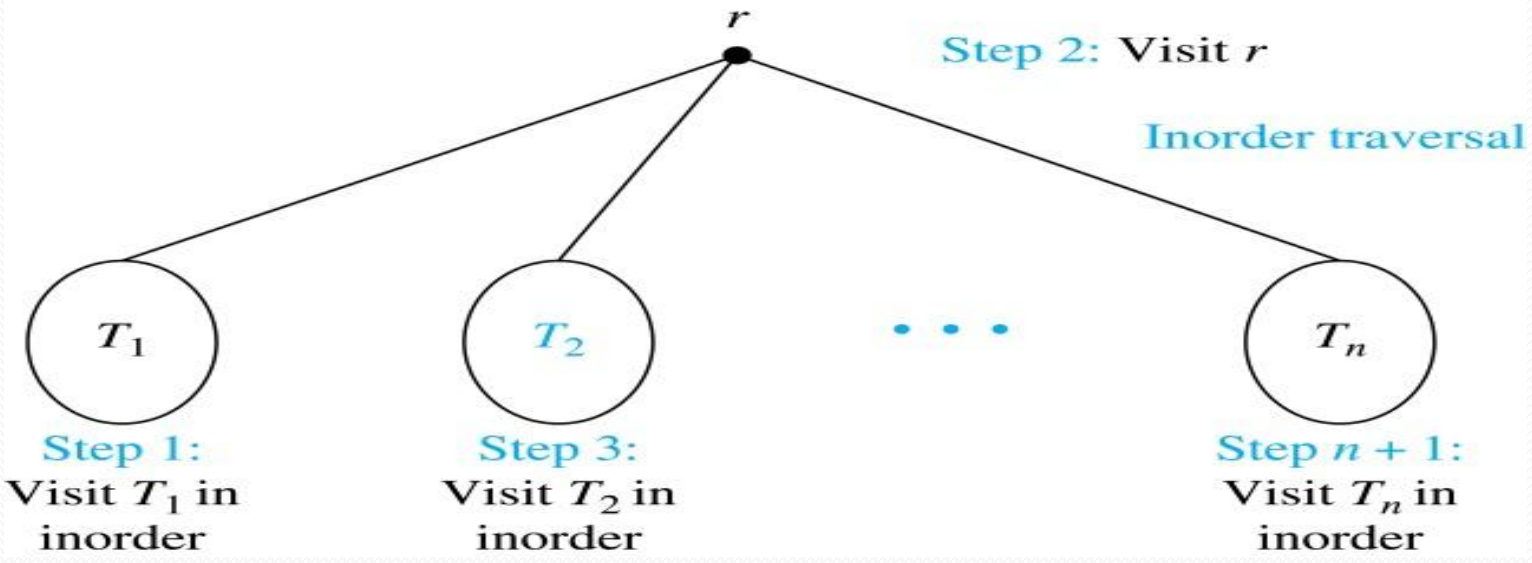






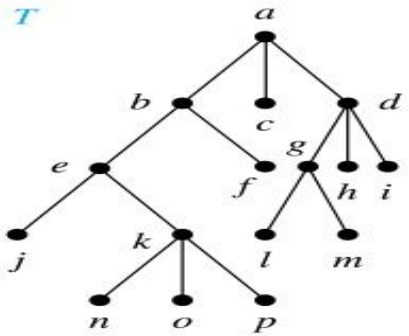
Inorder Traversal

Definition: Let T be an ordered rooted tree with root r. If T consists only of r, then r is the *inorder traversal* of T. Otherwise, suppose that T_1, T_2, ..., T_n are the subtrees of r from left to right in T. The inorder traversal begins by traversing T_1 in inorder, then visiting r, and continues by traversing T_2 in inorder, and so on, until T_n is traversed in inorder.

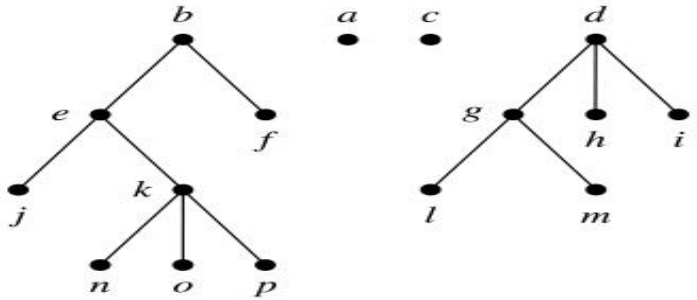


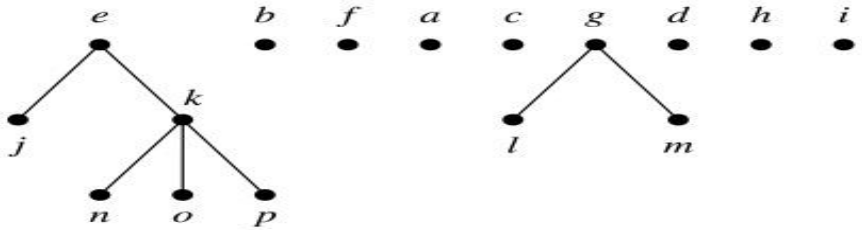
Inorder Traversal (*continued*)

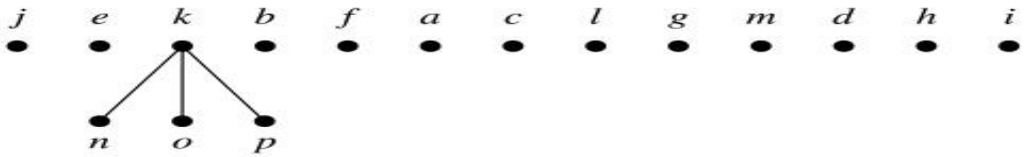
procedure
inorder (*T*: ordered rooted tree)
r := root of *T*
if *r* is a leaf **then** list *r*
else
  l := first child of *r* from left to right
  T(*l*) := subtree with *l* as its root
  inorder(*T*(*l*))
  list(*r*)
  for each child *c* of *r* from left to right
    T(*c*) := subtree with *c* as root
    inorder(*T*(*c*))



Inorder traversal: Visit leftmost subtree, visit root, visit other subtrees left to right

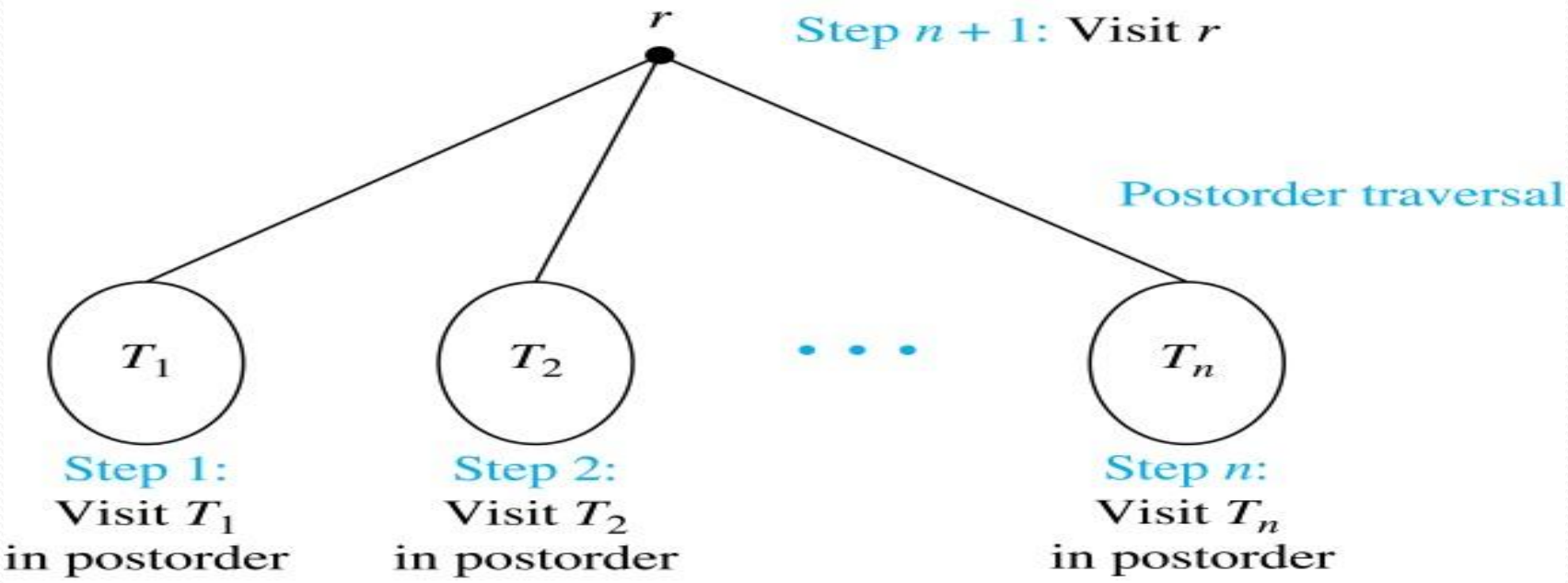






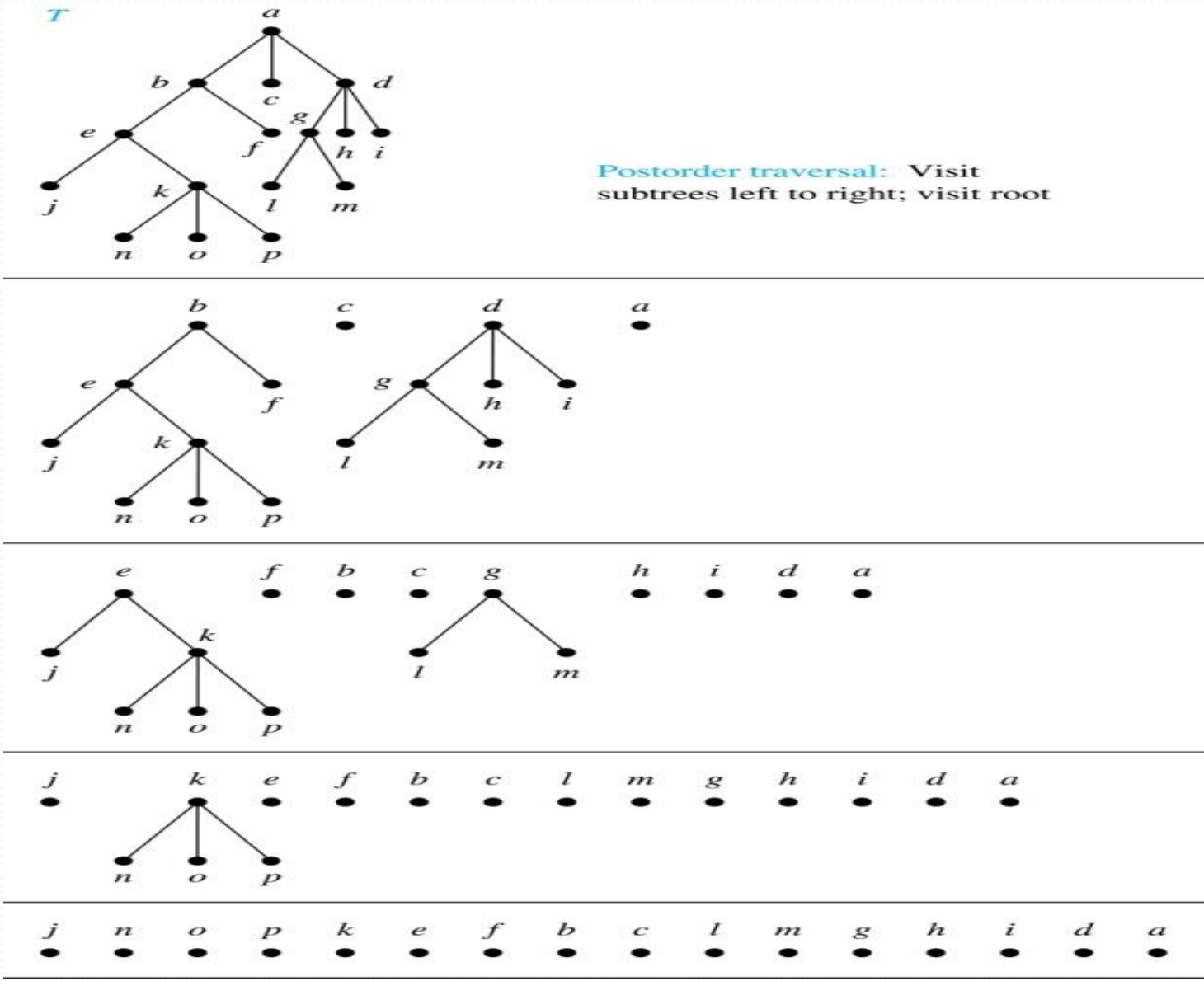
Postorder Traversal

Definition: Let T be an ordered rooted tree with root r. If T consists only of r, then r is the *postorder traversal* of T. Otherwise, suppose that T_1, T_2, ..., T_n are the subtrees of r from left to right in T. The postorder traversal begins by traversing T_1 in postorder, then T_2 in postorder, and so on, after T_n is traversed in postorder, r is visited.



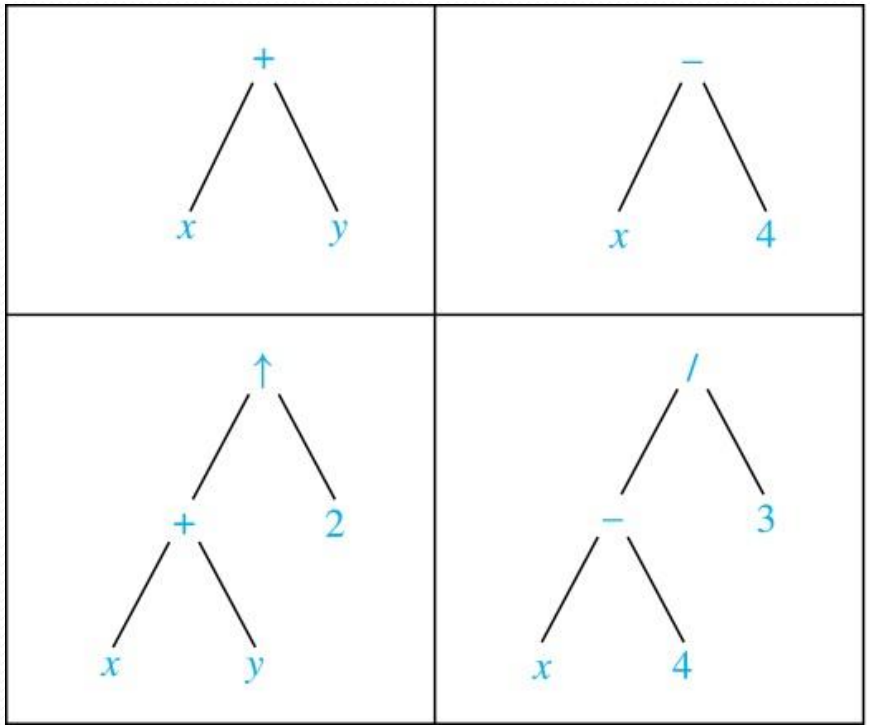
Post order Traversal (*continued*)

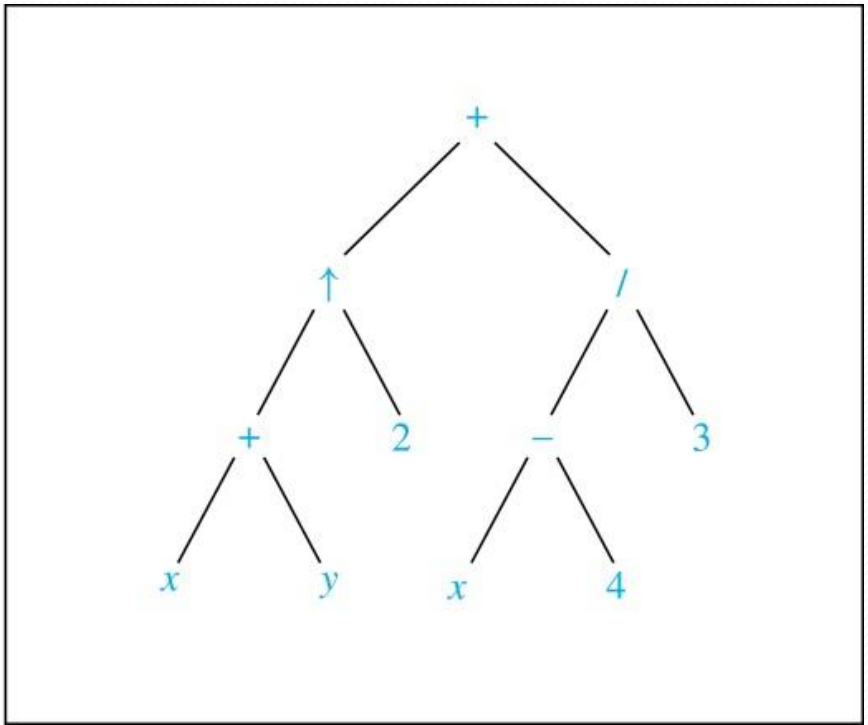
procedure *postordered* (*T*: ordered rooted tree)
r := root of *T*
for each child *c* of *r* from left to right
    T(*c*) := subtree with *c* as root
    postorder(*T*(*c*))
list *r*



Expression Trees

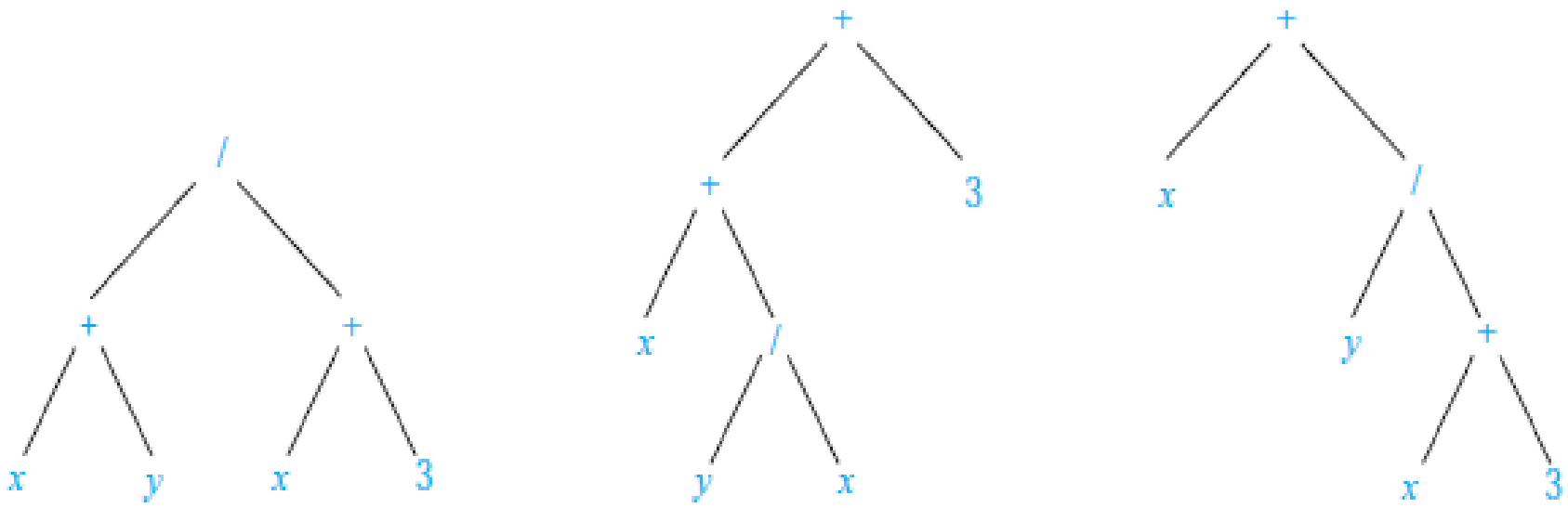
- Complex expressions can be represented using ordered rooted trees.
- Consider the expression $((x + y) \uparrow 2) + ((x - 4)/3)$.
- A binary tree for the expression can be built from the bottom up, as is illustrated here.



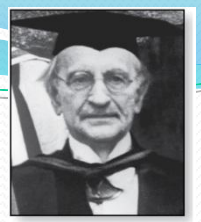


Infix Notation

- An inorder traversal of the tree representing an expression produces the original expression when parentheses are included except for unary operations, which now immediately follow their operands.
- We illustrate why parentheses are needed with an example that displays three trees all yield the same infix representation.



Rooted Trees Representing $(x + y)/(x + 3)$, $(x + (y/x)) + 3$, and $x + (y/(x + 3))$.



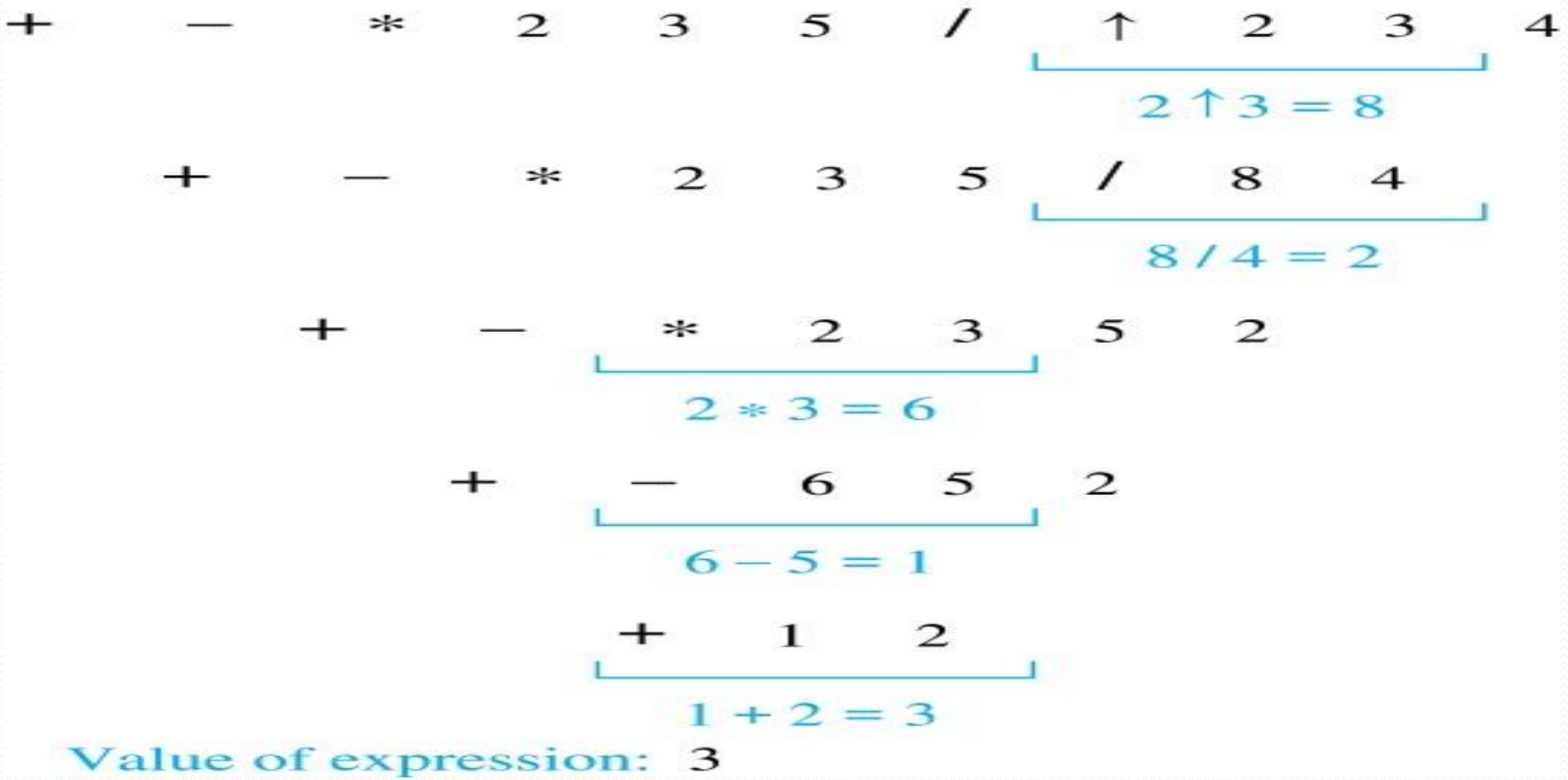
Jan Łukasiewicz
(1878-1956)

Prefix Notation

- When we traverse the rooted tree representation of an expression in preorder, we obtain the *prefix* form of the expression. Expressions in prefix form are said to be in *Polish notation*, named after the Polish logician Jan Łukasiewicz.
- Operators precede their operands in the prefix form of an expression. Parentheses are not needed as the representation is unambiguous.
- The prefix form of $((x + y) \uparrow 2) + ((x - 4)/3)$
  is $+ \uparrow + x\ y\ 2\ /\ -\ x\ 4\ 3$.
- Prefix expressions are evaluated by working from right to left. When we encounter an operator, we perform the corresponding operation with the two operations to the right.

Prefix Notation

- **Example**: We show the steps used to evaluate a particular prefix expression:

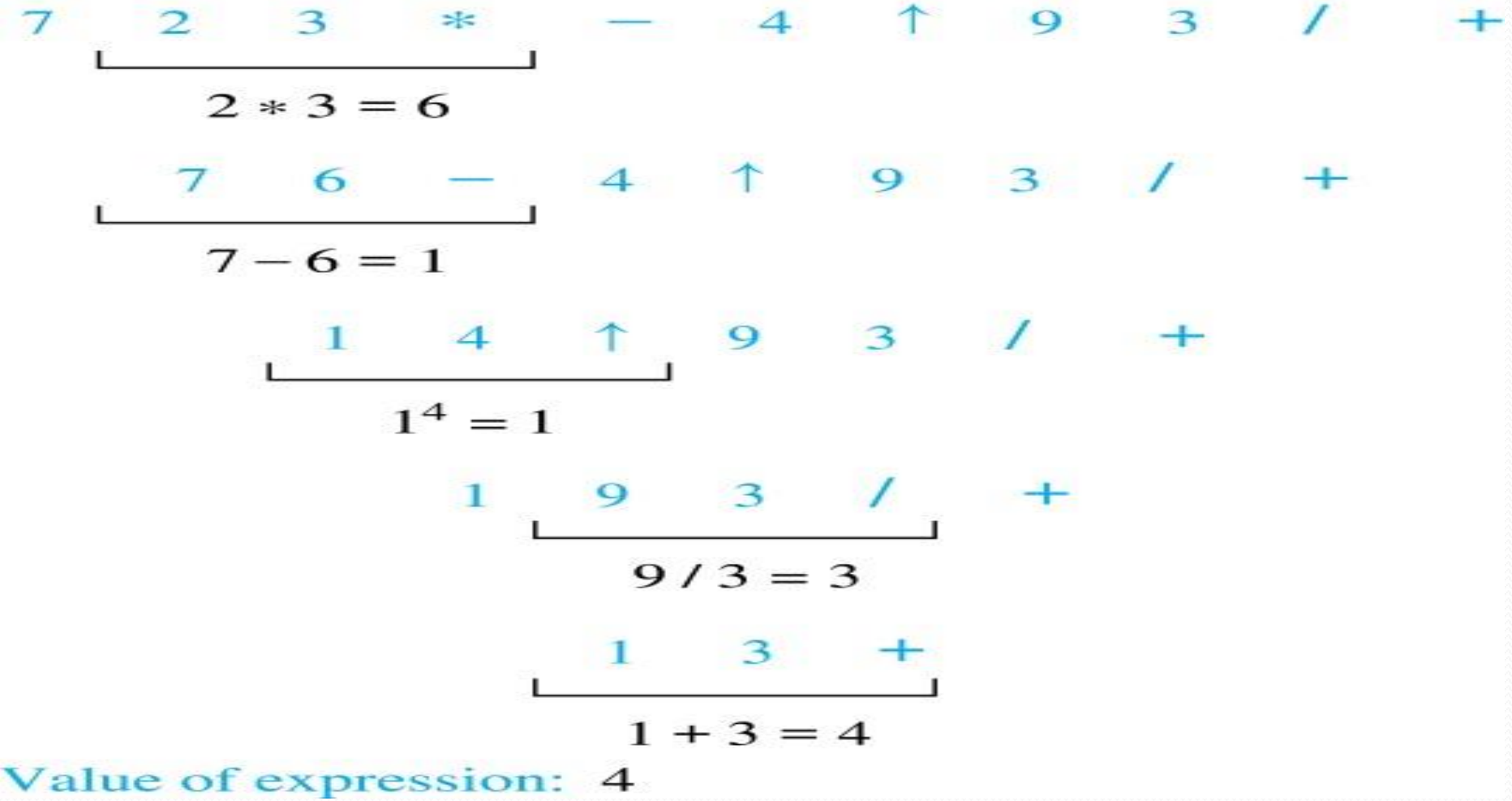


Postfix Notation

- We obtain the *postfix form* of an expression by traversing its binary trees in postorder. Expressions written in postfix form are said to be in *reverse Polish notation*.
- Parentheses are not needed as the postfix form is unambiguous.
- $x\ y + 2 \uparrow x\ 4 - 3 / +$ is the postfix form of $((x + y) \uparrow 2) + ((x - 4)/3)$.
- A binary operator follows its two operands. So, to evaluate an expression one works from left to right, carrying out an operation represented by an operator on its preceding operands.

Postfix Notation

- **Example**: We show the steps used to evaluate a particular postfix expression.



Spanning Trees

Section 11.4

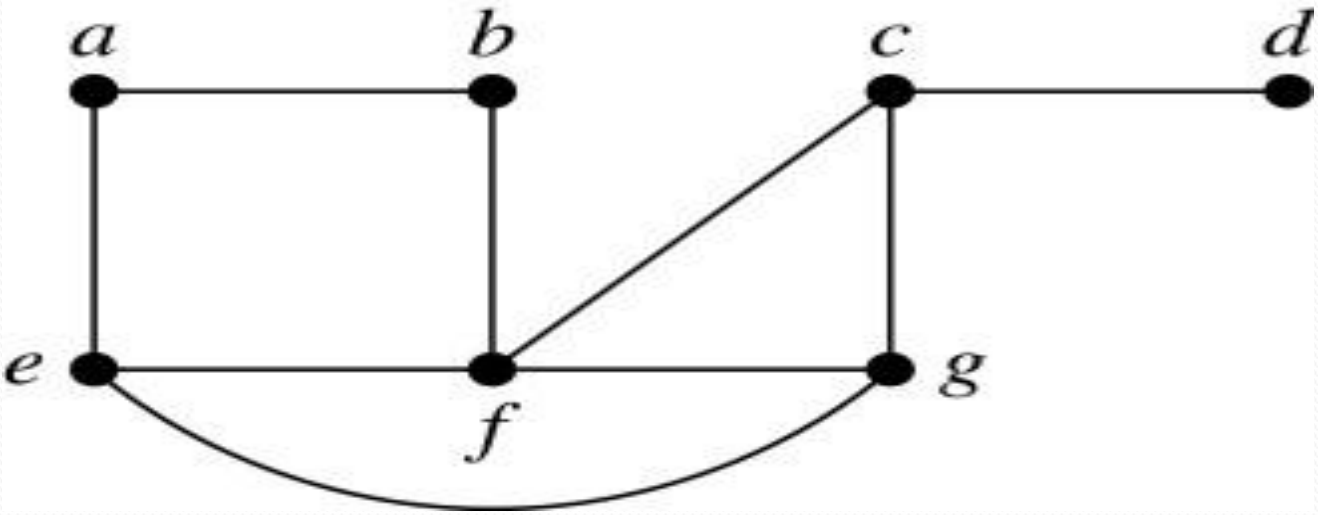
Section Summary

- Spanning Trees
- Prim's Algorithm
- Kruskal Algorithm

Spanning Trees

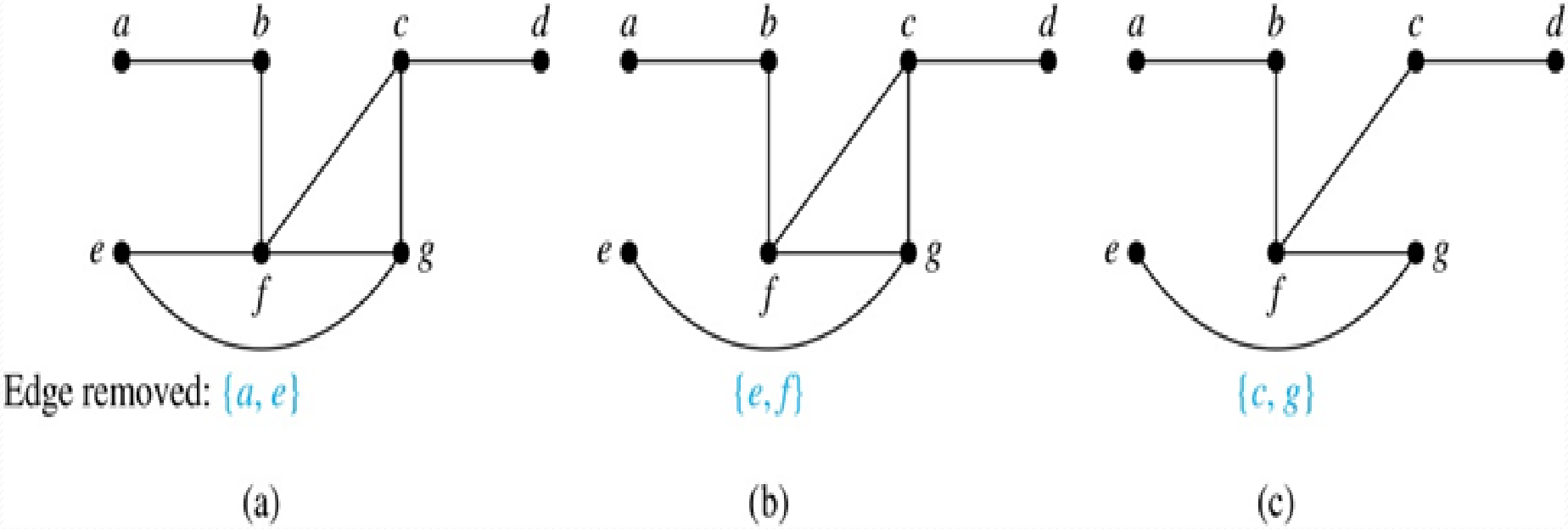
Definition: Let G be a simple graph. A spanning tree of G is a subgraph of G that is a tree containing every vertex of G.

Example: Find the spanning tree of the simple graph:



Spanning Trees

Solution: The graph is connected, but is not a tree because it contains simple circuits. Remove the edge $\{a, e\}$. Now one simple circuit is gone, but the remaining subgraph still has a simple circuit. Remove the edge $\{e, f\}$ and then the edge $\{c, g\}$ to produce a simple graph with no simple circuits. It is a spanning tree, because it contains every vertex of the original graph.



Minimum Spanning

Section 11.5

Minimum spanning tree

- A *minimum spanning tree* in a connected weighted graph is a spanning tree that has the smallest possible sum of weights of its edges.
- **Example:** A company plans to build a communications network connecting its five computer centers. Any pair of these centers can be linked with a leased telephone line. Which links should be made to ensure that there is a path between any two computer centers so that the total cost of the network is minimized?

Minimum spanning tree

- **Solution:** We can model this problem using the weighted graph shown in Figure 1, where vertices represent computer centers, edges represent possible leased lines, and the weights on edges are the monthly lease rates of the lines represented by the edges. We can solve this problem by finding a spanning tree so that the sum of the weights of the edges of the tree is minimized. Such a spanning tree is called a **minimum spanning tree**.

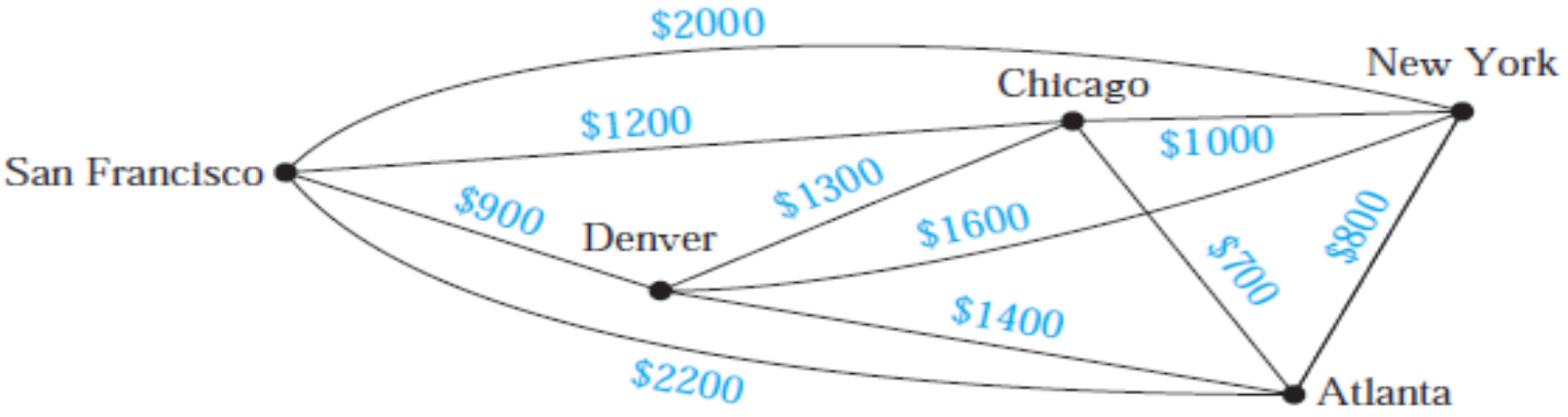


FIGURE 1 A Weighted Graph Showing Monthly Lease Costs for Lines in a Computer Network.

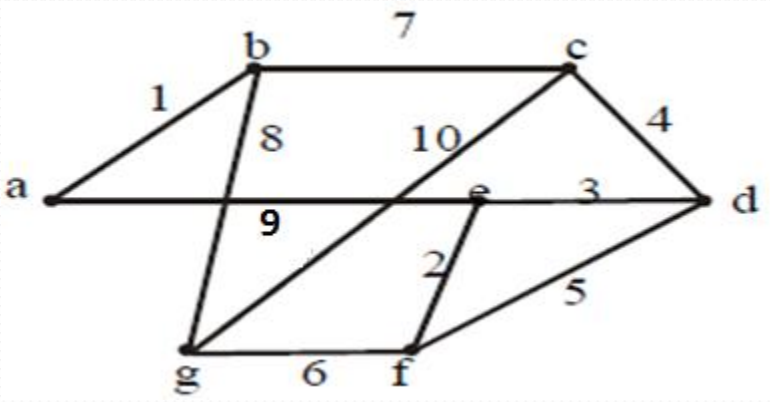
PRIM'S ALGORITHM

ALGORITHM 1 Prim's Algorithm.

procedure *Prim*(G: weighted connected undirected graph with n vertices)
T := a minimum-weight edge
for $i := 1$ **to** $n - 2$
    e := an edge of minimum weight incident to a vertex in T and not forming a simple circuit in T if added to T
    $T := T$ with e added
return T {T is a minimum spanning tree of G}

Minimal spanning tree (MST)

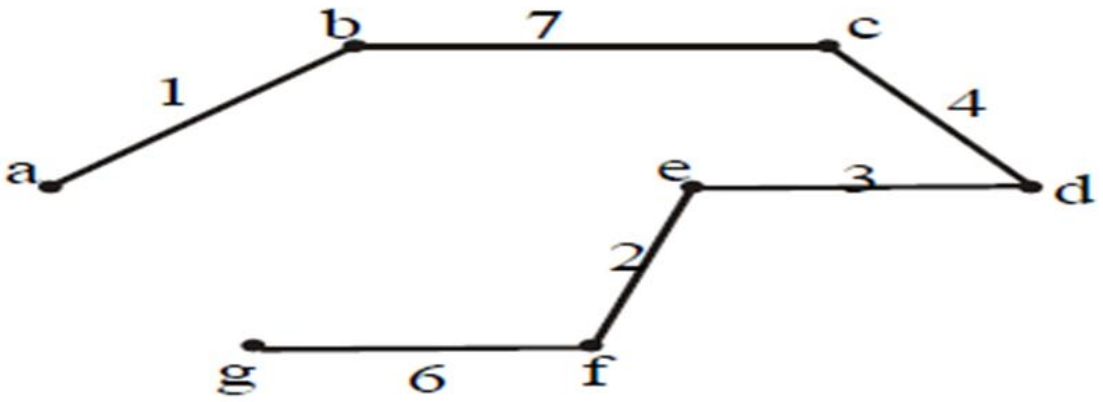
Example: Use Prims algorithm to find a minimal spanning tree for the graph below. Indicate the order in which edges are added to form the tree.



Order of adding the edges:

{a , b}, {b , c}, {c , d}, {d , e}, {e , f}, {f , g}

MST COST = 23



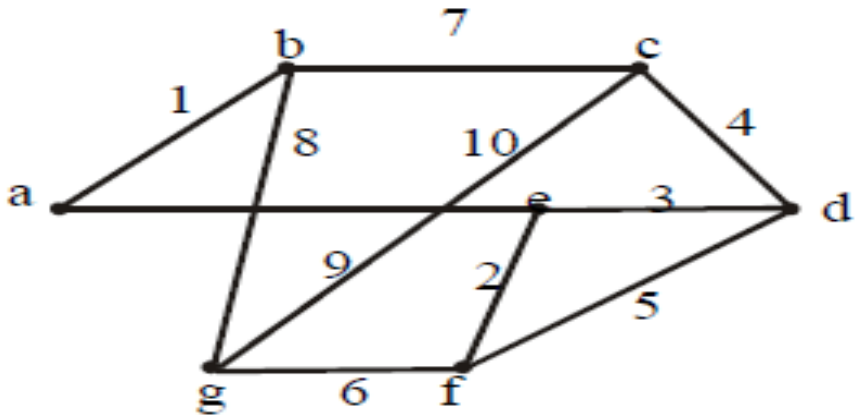
KRUSKAL'S ALGORITHM

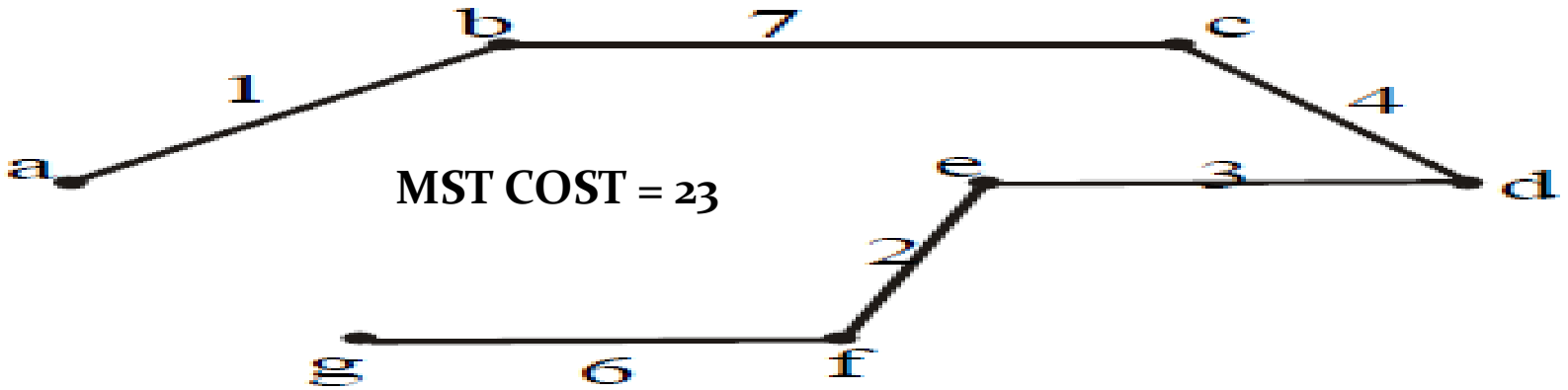
ALGORITHM 2 Kruskal's Algorithm.

procedure *Kruskal*(G: weighted connected undirected graph with n vertices)
T := empty graph
for $i := 1$ **to** $n - 1$
    e := any edge in G with smallest weight that does not form a simple circuit when added to T
    $T := T$ with e added
return T {T is a minimum spanning tree of G}

Minimal spanning tree (MST)

Example: Use Kruskal's algorithm to find a minimal spanning tree for the graph below. Indicate the order in which edges are added to form the tree.





Order of adding the edges:
{a, b}, {e, f}, {e, d}, {c, d}, {g, f}, {b, c}