GRAPH THEORY

An Algorithmic Approach

Nicos Christofides
Management Science
Imperial College
London

1975
ACADEMIC PRESS
London New York
San Francisco
A Subsidiary of Harcourt Brace Jovanovich, Publishers

Computer Science and Applied Mathematics

A SERIES OF MONOGRAPHS AND TEXTBOOKS

Editor
Werner Rheinboldt
University of Maryland

HANS P. KÜNZI, H. G. TZSCHACH, and C. A. ZEHNDER. Numerical Methods of Mathematical Optimization: With ALGOL and FORTRAN Programs, Corrected and Augmented Edition

AZRIEL ROSENFELD. Picture Processing by Computer

JAMES ORTEGA AND WERNER RHEINBOLDT. Iterative Solution of Nonlinear Equations in Several Variables

AZARIA PAZ. Introduction to Probabilistic Automata

DAVID YOUNG. Iterative Solution of Large Linear Systems

ANN YASUHARA. Recursive Function Theory and Logic

JAMES M. ORTEGA. Numerical Analysis: A Second Course

G. W. STEWART. Introduction to Matrix Computations

CHIN-LIANG CHANG AND RICHARD CHAR-TUNG LEE. Symbolic Logic and Mechanical Theorem Proving

C. C. GOTLIEB AND A. BORODIN. Social Issues in Computing

ERWIN ENGELER. Introduction to the Theory of Computation

F. W. J. OLVER. Asymptotics and Special Functions

DIONYSIOS C. TSICHRITZIS AND PHILIP A. BERNSTEIN. Operating Systems

ROBERT R. KORFHAGE. Discrete Computational Structures

PHILIP J. DAVIS AND PHILIP RABINOWITZ. Methods of Numerical Integration

A. T. BERZTISS. Data Structures: Theory and Practice, Second Edition (in preparation)

N. CHRISTOFIDES. Graph Theory: An Algorithmic Approach

In preparation

ALBERT NIJENHUIS AND H. S. WILF. Combinatorial Algorithms

ACADEMIC PRESS INC. (LONDON) LTD.
24/28 Oval Road.
London NW1

United States Edition published by
ACADEMIC PRESS INC.
111 Fifth Avenue
New York, New York 10003

Second printing 1977

Library of Congress Catalog Card Number: 74–5664
ISBN: 0-12-174350-0

Set in 'Monophoto' Times and printed offset litho in Great Britain by
Page Bros (Norwich) Ltd. Norwich

Preface

It is often helpful and visually appealing, to depict some situation which is of interest by a graphical figure consisting of points (vertices)—representing entities—and lines (links) joining certain pairs of these vertices and representing relationships between them. Such figures are known by the general name *graphs* and this book is devoted to their study. Graphs are met with everywhere under different names: "structures" in civil engineering, "networks" in electrical engineering, "sociograms", "communication structures" and "organizational structures" in sociology and economics, "molecular structure" in chemistry, "road maps", gas or electricity "distribution networks" and so on.

Because of its wide applicability, the study of graph theory has been expanding at a very rapid rate during recent years; a major factor in this growth being the development of large and fast computing machines. The direct and detailed representation of practical systems, such as distribution or telecommunication networks, leads to graphs of large size whose successful analysis depends as much on the existence of "good" algorithms as on the availability of fast computers. In view of this, the present book concentrates on the development and exposition of algorithms for the analysis of graphs, although frequent mention of application areas is made in order to keep the text as closely related to practical problem-solving as possible. By so doing, it is hoped that the reader will be left in a position to relate and adapt the basic concepts to his own field of application, and, indeed, be able to derive new methods of solution to his specific problem.

Although, in general, algorithmic efficiency is considered of prime importance, the present text is not meant to be a handbook of efficient algorithms. Often a method is discussed because of its close relation to (or derivation from) previously introduced concepts, in preference to another algorithm which may be equally—and in some cases slightly more—efficient. The overriding consideration is to leave the reader with as coherent a body of knowledge with regard to graph analysis algorithms, as possible.

The title *Graph Theory* must, to some extent, be a misnomer on any single volume, since it is quite impossible to cover even remotely the subject in such a short space. The present book is no exception and its contents reflect, as they must, the author's interest and background in Operations Research, Computer and Management Science.

Chapter 2 discusses basic reachability and connectivity properties of graphs, the computation of strong components and bases and their application to the formation of power-groups and coalitions in organizations.

Chapter 3 considers two problems in connection with choosing extremal subsets of vertices with prescribed properties. The problems of computing maximal independent or minimal dominating sets are discussed, the latter problem generalizing to the set covering problem. Applications of the set covering problem to airline crew scheduling, vehicle scheduling and information retrieval are given and the transformation of other graph theory problems to the set covering problem discussed.

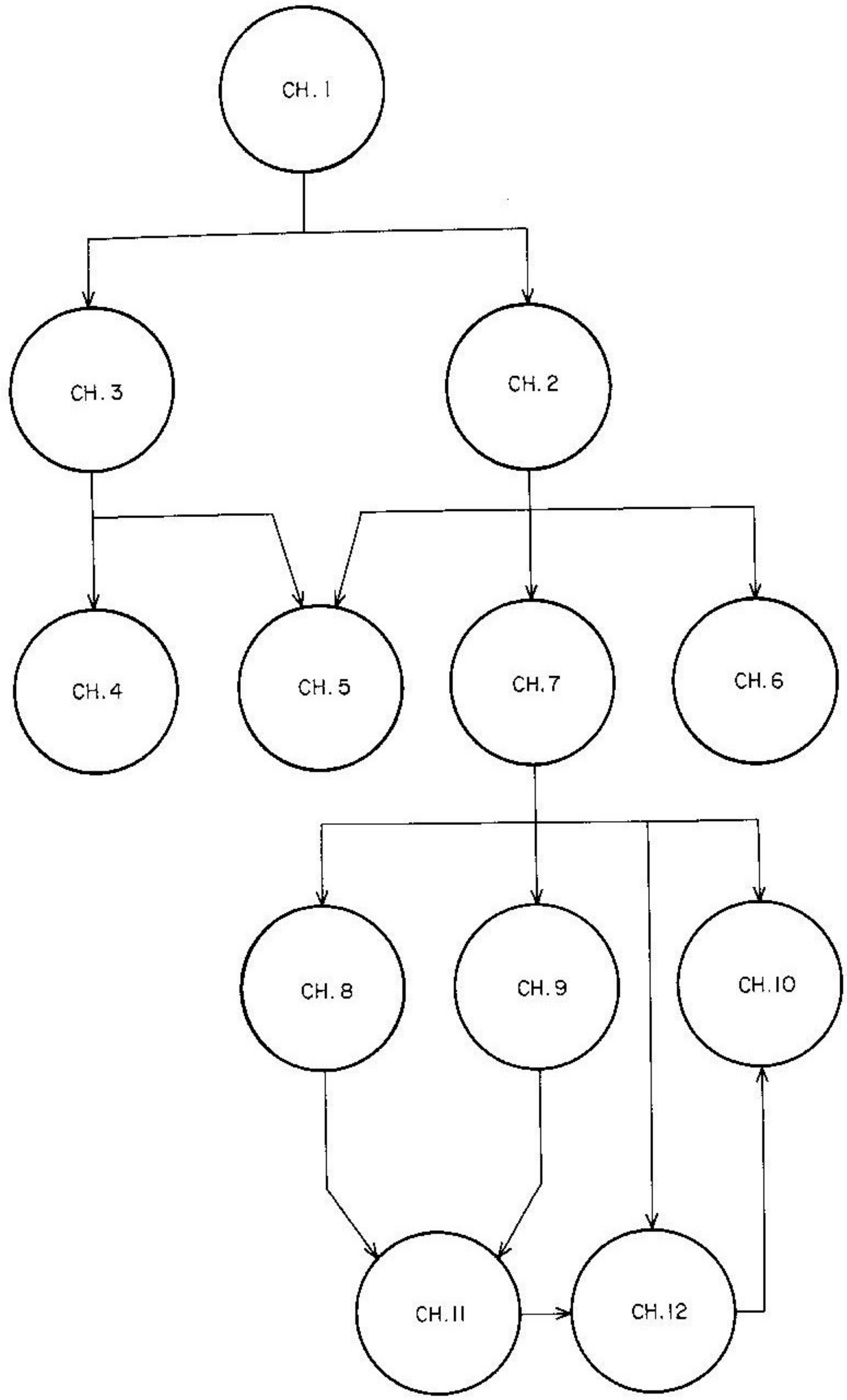
Chapter 4 considers the vertex colouring problem, a specific case of the set covering problem discussed in the previous Chapter. Applications of the colouring problem to the scheduling of timetables and resource allocation are given, together with generalizations to the loading problem.

The next two Chapters are concerned with location problems on graphs. Chapter 5 considers the problem of locating multi-centres on a graph and is applicable to cases of locating emergency facilities such as fire, police or ambulance stations on a road network. Chapter 6 considers the problem of locating multi-medians on a graph and is applicable to cases of locating facilities such as depots or switching centres in goods-distribution or telecommunication networks. Chapter 5 deals with a minimax and Chapter 6 with a minisum location problem.

Chapter 7 is concerned with trees, shortest spanning trees and the Steiner problem, and their application to the construction of electric power or gas pipeline networks.

Chapter 8 deals with the problem of finding shortest paths between pairs of vertices and its applications to routing problems for maximizing capacity or reliability, and to the special case of PERT networks and critical path analysis. The Chapter also deals with the determination of negative cost circuits in graphs and its applications to other graph theoretic problems. The calculation of second, third, etc. shortest paths is also discussed.

The next two Chapters deal with problems of finding circuits in graphs. Chapter 9 discusses general circuits and cut-sets. The Chapter also deals with Eulerian circuits and the Chinese postman's problem with its applications to refuse collection, delivery of milk or post and inspection of distributed parameter systems. Chapter 10 is in two parts. The first part deals with the problem of finding a Hamiltonian circuit in a graph that is not complete and its application to machine scheduling. The second part of this Chapter deals with the problem of finding the shortest Hamiltonian circuit i.e. with the well known "travelling salesman problem" and its applications to vehicle routing.



Main relationships between chapters of this book

Chapter 11 is concerned with maximum flow and minimum cost—maximum flow problems in graphs with arc capacities and costs. Flow problems in graphs with arc gains occur in mathematical models of arbitrage, current flow in active electric circuits, etc., and are also considered.

Chapter 12 deals with the problem of finding maximum matchings in graphs and describes the generalized hungarian algorithm. The algorithm particularizes to the bipartite graphs case for the assignment and transportation problems, both of which are of great significance in Operations Research with applications to assignment of people to jobs, facilities to locations, aircraft to routes and so on.

The relationships between the Chapters of this book are portrayed by the graph on page vii.

Parts of the contents of the book were developed with the help of grants from the Science Research Council, for research in mathematical programming. This assistance is gratefully acknowledged. In preparing this book I have benefited from the help of many people. I want especially to thank Professor Sam Eilon, head of the Department of Management Science at Imperial College, Peter Viola, Geff Selby, Peter Brooker and Sam Korman. I also wish to thank Professor Donald Knuth who has read an earlier version of the manuscript and pointed out several errors. For the arduous task of typing the manuscript I am indebted to the skill and patience of Miss Margaret Hudgell. Finally, I must mention the invaluable help of my wife Ann, who not only gave up countless evenings during the writing of this book but also for her assistance with the proof reading.

Nicos Christofides
June 1975
London.

List of Symbols

$\mathbf{A} = [a_{ij}]$	Adjacency matrix
A	Set of arcs
a_j	j^{th} arc
$\mathbf{B} = [b_{ij}]$	Incidence matrix
$\mathbf{C} = [c_{ij}]$	Arc cost matrix
$c_{ij} = c(x_i, x_j)$	Cost of arc (x_i, x_j)
c_j	Cost of arc a_j
$\mathbf{D} = [d_{ij}]$	Shortest distance matrix
$d_{ij} = d(x_i, x_j)$	Shortest distance (cost of least cost path) from x_i to x_j
$d_i = d(x_i)$	Degree of vertex x_i
$d_i^H = d^H(x_i)$	Degree of vertex x_i with respect to graph H
$d_0(x_i)$, $d_t(x_i)$	Outdegree and indegree of vertex x_i respectively
E	Covering
f_{ij}	Value of the maximum flow from x_i to x_j
$G = (X, A)$	Graph with set of vertices X and set of arcs A
$G(\xi)$	Graph G with flow pattern ξ flowing in it
$G^\mu(\xi)$	Incremental graph of $G(\xi)$
$\bar{G}$	Graph G with all arc directions ignored
$\tilde{G}$	Graph complementary to G
g_{ij}	Gain of arc (x_i, x_j)
$\mathbf{K} = [k_{ij}]$	Cut-set matrix
K	Cut-set
m	Number of arcs in graph
M	Matching
n	Number of vertices in graph
$p(x_i)$	Predecessor to vertex x_i
$\mathbf{Q}$	Reaching matrix
$Q(x_i)$	Reaching set of vertex x_i
$q_{ij} = q(x_i, x_j)$	Upper limit (capacity) on the flow in arc (x_i, x_j)
$\mathbf{R}$	Reachability matrix
$R(x_i)$	Reachability set of vertex x_i
r_{ij}	Lower limit on the flow in arc (x_i, x_j)
s	Initial vertex in shortest path or flow calculation
t	Final vertex in shortest path or flow calculation
v_i	"Weight" of vertex x_i
x_i	i^{th} vertex
X	Set of vertices of graph G
$\alpha[G]$	Independence number
$\beta[G]$	Dominance number
$\gamma[G]$	Chromatic number

Γ	Set of correspondences
$\Gamma(x_i)$	Set of vertices x_j such that $(x_i, x_j) \in A$
$\Gamma^{-1}(x_i)$	Set of vertices x_j such that $(x_j, x_i) \in A$
$\Theta = [\theta_{ij}]$	Matrix storing shortest paths
$\xi_{ij} = \xi(x_i, x_j)$	Flow in arc (x_i, x_j)
ξ_{ij}^e, ξ_{ij}^0	Entering and leaving flow in arc (x_i, x_j) respectively
$\hat{\Xi}$	Maximum flow pattern
$\tilde{\Xi}$	Optimum flow pattern
$\hat{\tilde{\Xi}}$	Optimum maximum flow pattern
$\mathbf{\Phi} = [\phi_{ij}]$	Circuit matrix
Φ	Circuit

Contents

Chapter 3. Independent and Dominating Sets—The Set Covering Problem

Chapter 4. Graph Colouring

Chapter 5. The Location of Centres

Chapter 6. The Location of Medians

Chapter 7. Trees

Chapter 8. Shortest Paths

Chapter 9. Circuits, Cut-sets and Euler's Problem

Chapter 10. Hamiltonian Circuits, Paths and the Travelling Salesman Problem

Part I

Part II

Chapter 11. Network Flows

Chapter 12. Matchings, Transportation and Assignment Problems

Chapter 1

Introduction

1. Graphs—Definition

A *graph* G is a collection of points or *vertices* $x_1, x_2, \ldots, x_n$ (denoted by the set X), and a collection of lines $a_1, a_2, \ldots, a_m$ (denoted by the set A) joining all or some of these points. The graph G is then fully described and denoted by the doublet (X, A).

If the lines in A have a direction—which is usually shown by an arrow—they are called *arcs* and the resulting graph is called a *directed* graph (Fig. 1.1(a)). If the lines have no orientation they are called *links* and the graph is *nondirected* (Fig. 1.1(b)). In the case where $G = (X, A)$ is a directed graph but we want to disregard the direction of the arcs in A, the nondirected counterpart to G will be written as $\bar{G} = (X, \bar{A})$.

Throughout this book, when an arc is denoted by the pair of its *initial* and *final* vertices (i.e. by its two *terminal* vertices), its direction will be assumed to be from the first vertex to the second. Thus, in Fig. 1.1(a) (x_1, x_2) refers to arc a_1, and (x_2, x_1) to arc a_2.

An alternative and often preferable way to describe a directed graph G, is by specifying the set X of vertices and a *correspondence* Γ which shows how the vertices are related to each other. Γ is called a mapping of the set X in X and the graph is denoted by the doublet $G = (X, \Gamma)$.

In the example of Fig. 1.1(a) we have

$\Gamma(x_1) = \{x_2, x_5\}$ i.e. x_2 and x_5 are the final vertices of arcs whose initial vertex is x_1.

$$\Gamma(x_2) = \{x_1, x_3\}$$

$$\Gamma(x_3) = \{x_1\}$$

$\Gamma(x_4) = \varnothing$, the *null* or empty set

and

$$\Gamma(x_5) = \{x_4\}.$$

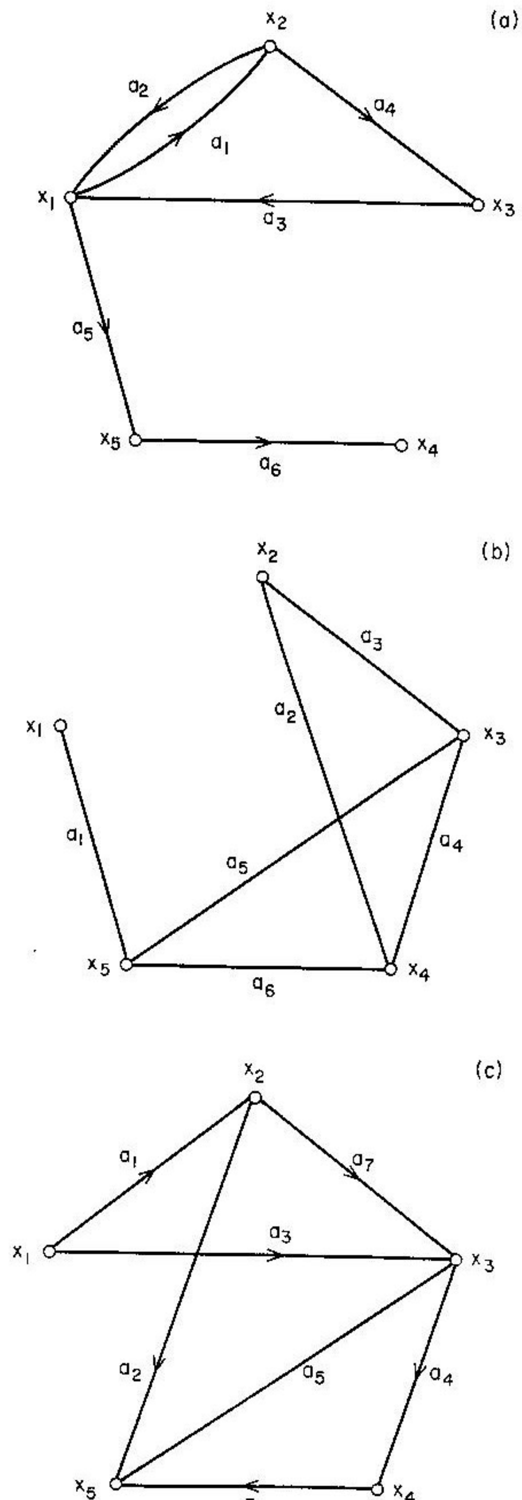


FIG. 1.1. (a) Directed, (b) Nondirected and (c) Mixed graphs

In the case of a nondirected graph, or of a graph containing both arcs and links (such as the graphs shown in Figs 1.1(b) and 1.1(a)), the correspondences Γ will be assumed to be those of an equivalent directed graph in which every link has been replaced by two arcs in opposite directions. Thus, for the graph of Fig. 1.1(b) for example, $\Gamma(x_5) = \{x_1, x_3, x_4\}$, $\Gamma(x_1) = \{x_5\}$ etc.

Since $\Gamma(x_i)$ has been defined to mean the set of those vertices $x_j \in X$ for which an arc (x_i, x_j) exists in the graph, it is natural to write as $\Gamma^{-1}(x_i)$ the set of those vertices x_k for which an arc (x_k, x_i) exists in G. The relation $\Gamma^{-1}(x_i)$ is then called the *inverse correspondence*. Thus for the graph shown in Fig. 1.1(a) we have:

$$\Gamma^{-1}(x_1) = \{x_2, x_3\}$$

$$\Gamma^{-1}(x_2) = \{x_1\}$$

etc.

It is quite obvious that for a nondirected graph $\Gamma^{-1}(x_i) = \Gamma(x_i)$ for all $x_i \in X$.

When the correspondence Γ does not operate on a single vertex but on a set of vertices such as $X_q = \{x_1, x_2, \ldots, x_q\}$, then $\Gamma(X_q)$ is taken to mean:

$$\Gamma(X_q) = \Gamma(x_1) \cup \Gamma(x_2) \cup \ldots\ldots \cup \Gamma(x_q)$$

i.e. $\Gamma(X_q)$ is the set of those vertices $x_j \in X$ for which at least one arc (x_i, x_j) exists in G, for some $x_i \in X_q$. Thus, for the graph of Fig. 1.1(a); $\Gamma(\{x_2, x_5\}) = \{x_1, x_3, x_4\}$ and $\Gamma(\{x_1, x_3\}) = \{x_2, x_5, x_1\}$.

The double correspondence $\Gamma(\Gamma(x_i))$ is written as $\Gamma^2(x_i)$. Similarly the triple correspondence $\Gamma(\Gamma(\Gamma(x_i)))$ is written as $\Gamma^3(x_i)$ and so on. Thus, the graph in Fig. 1.1(a):

$$\Gamma^2(x_1) = \Gamma(\Gamma(x_1)) = \Gamma(\{x_2, x_5\}) = \{x_1, x_3, x_4\}$$

$$\Gamma^3(x_1) = \Gamma(\Gamma^2(x_1)) = \Gamma(\{x_1, x_3, x_4\}) = \{x_1, x_2, x_5\}$$

etc.

Similarly for $\Gamma^{-2}(x_i)$, $\Gamma^{-3}(x_i)$ and so on.

2. Paths and Chains

A *path* in a directed graph is any sequence of arcs where the final vertex of one is the initial vertex of the next one.

Thus in Fig. 1.2 the sequence of arcs:

$$a_6, a_5, a_9, a_8, a_4 \tag{1.1}$$

$$a_1, a_6, a_5, a_9 \tag{1.2}$$

$$a_1, a_6, a_5, a_9, a_{10}, a_6, a_4 \tag{1.3}$$

are all paths.

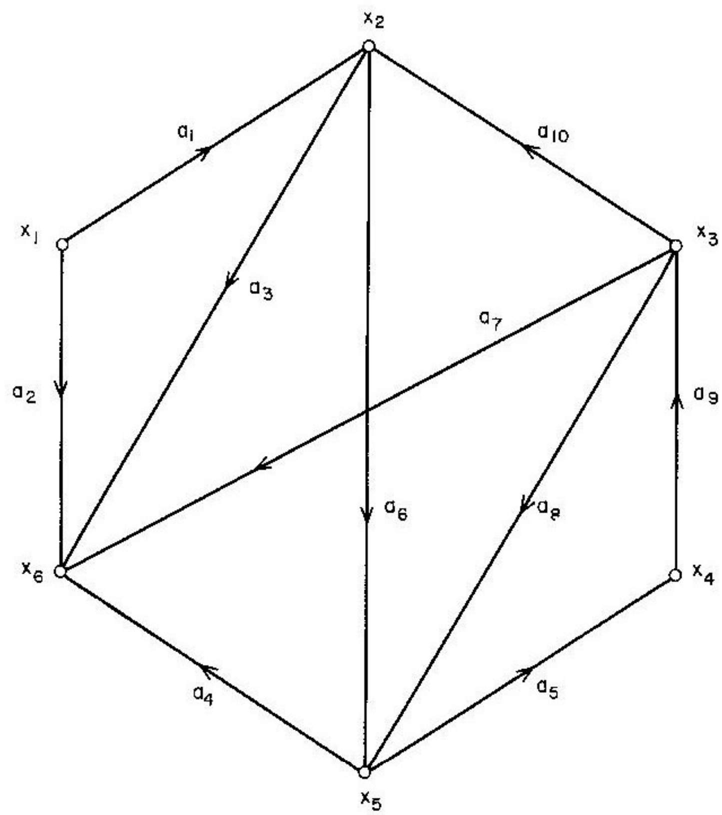


FIG. 1.2

Arcs $a = (x_i, x_j)$, $x_i \neq x_j$ which have a common terminal vertex are called *adjacent*. Also, two vertices x_i and x_j are called adjacent if either arc (x_i, x_j) or arc (x_j, x_i) or both exist in the graph. Thus in Fig. 1.2 arcs a_1, a_{10}, a_3 and a_6 are adjacent and so are the vertices x_5 and x_3; on the other hand arcs a_1 and a_5 or vertices x_1 and x_4 are not adjacent.

A *simple* path is a path which does not use the same arc more than once. Thus the paths (1.1) and (1.2) above are simple but path (1.3) is not, since it uses arc a_6 twice.

An *elementary* path is a path which does not use the same vertex more than once. Thus the path (1.2) is elementary but paths (1.1) and (1.3) are not. Obviously an elementary path is also simple but the reverse is not necessarily true. Note for example that path (1.1) is simple but not elementary, that path (1.2) is both simple and elementary and that path (1.3) is not simple and not elementary.

A *chain* is the nondirected counterpart of the path and applies to graphs with the direction of its arcs disregarded. Thus a chain is a sequence of links

$(\bar{a}_1, \bar{a}_2, \ldots, \bar{a}_q)$ in which every link $\bar{a}_i$, except perhaps the first and last links, is connected to the links $\bar{a}_{i-1}$ and $\bar{a}_{i+1}$ by its two terminal vertices.

$$\bar{a}_2, \bar{a}_4, \bar{a}_8, \bar{a}_{10} \tag{1.4}$$

$$\bar{a}_2, \bar{a}_7, \bar{a}_8, \bar{a}_4, \bar{a}_3 \tag{1.5}$$

and

$$\bar{a}_{10}, \bar{a}_7, \bar{a}_4, \bar{a}_8, \bar{a}_7, \bar{a}_2 \tag{1.6}$$

are all chains; where a bar above the symbol of an arc means that its direction is disregarded i.e. it is to be considered as a link.

In an exactly analogous way as we defined simple and elementary paths we can define simple and elementary chains. Thus, chain (1–4) is simple and elementary, chain (1–5) is simple but not elementary and chain (1–6) is both not simple and not elementary.

A path or a chain may also be represented by the sequence of vertices that are used. Thus, path (1.1) may also be represented by the sequence $x_2, x_5, x_4, x_3, x_5, x_6$ of vertices, and this representation is often more useful when one is concerned with finding elementary paths or chains.

2.1 Weights and the length of a path

A number c_{ij} may sometimes be associated with an arc (x_i, x_j). These numbers are called *weights*, *lengths* or *costs* and the graph is then called *arc-weighted*. Also a weight v_i may sometimes be associated with a vertex x_i and the resulting graph is then called *vertex-weighted*. If a graph is both arc and vertex weighted it is simply called weighted.

Considering a path μ represented by the sequence of arcs $(a_1, a_2, \ldots, a_q)$, the *length* (*or cost*) *of the path* $l(\mu)$ is taken to be the sum of the arc weights on the arcs appearing in μ, i.e.

$$l(\mu) = \sum_{(x_i, x_j) \text{ in } \mu} c_{ij}$$

Thus, the words "length", "cost" and "weight" when applied to arcs, are all considered to be equivalent, and in specific applications that word will be chosen which gives the best intuitive meaning and which is in agreement with the usage found in the literature.

The *cardinality* of the path μ is q i.e. the number of arcs appearing in the path.

3. Loops, Circuits and Cycles

A *loop* is an arc whose initial and final vertices are the same. In Fig. 1.3 for example arcs a_2 and a_5 form loops.

A *circuit* is a path $a_1, a_2, \ldots, a_q$ in which the initial vertex of a_1 coincides with the final vertex of a_q.

Thus in Fig. 1.3 the sequences:

$$a_3, a_6, a_{11} \tag{1.7}$$

$$a_{11}, a_3, a_4, a_7, a_1, a_{12}, a_9 \tag{1.8}$$

$$a_3, a_4, a_7, a_{10}, a_9, a_{11} \tag{1.9}$$

are all circuits.

Circuits (1.7) and (1.9) are elementary since they do not use the same vertex more than once (except for the initial and final vertices which are the same), but circuit (1.8) is not elementary since vertex x_1 is used twice.

An elementary circuit which passes through all the n vertices of a graph G is of special significance and is known as a *Hamiltonian* circuit. Of course not all graphs have a Hamiltonian circuit. Thus, circuit (1.9) is a Hamiltonian circuit of the graph of Fig. 1.3; but the graph of Fig. 1.2 has no Hamiltonian circuits as can be seen quite easily from the fact that there is no arc having x_1 as its final vertex.

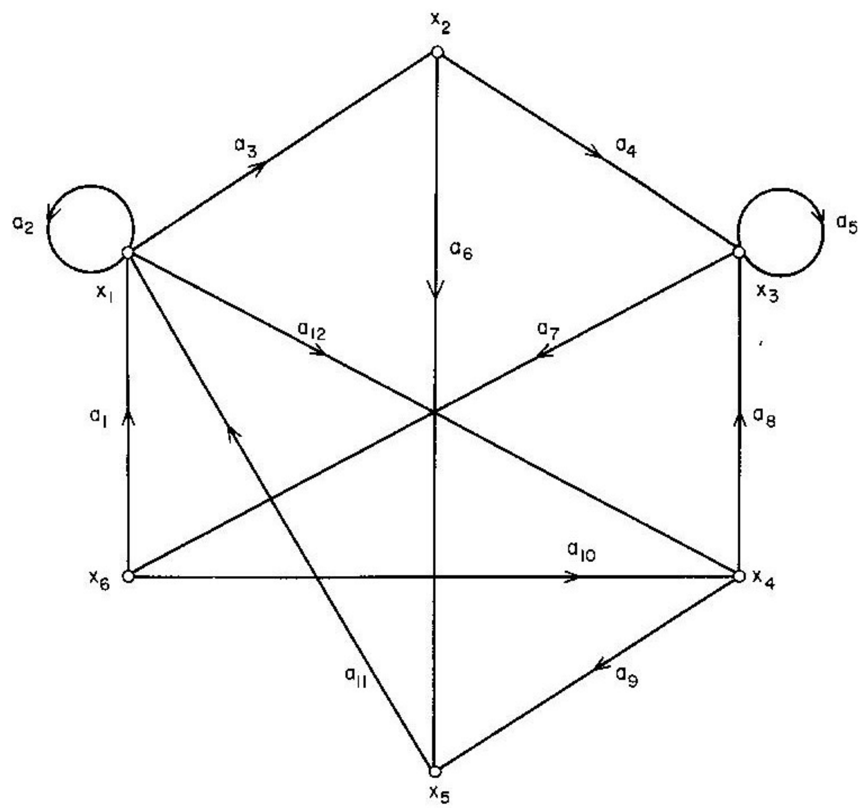


FIG. 1.3

A *cycle* is the nondirected counterpart of the circuit. Thus, a cycle is a chain $x_1, x_2, \ldots, x_q$ in which the beginning and end vertices are the same, i.e. in which $x_1 = x_q$.

In Fig. 1.3 the chains:

$$\bar{a}_1, \bar{a}_{12}, \bar{a}_{10} \tag{1.10}$$

and

$$\bar{a}_{10}, \bar{a}_1, \bar{a}_3, \bar{a}_4, \bar{a}_7, \bar{a}_1, \bar{a}_{12} \tag{1.11}$$

all form cycles.

4. Degrees of a Vertex

The number of arcs which have a vertex x_i as their initial vertex is called the *outdegree* of vertex x_i, and similarly the number of arcs which have x_i as their final vertex is called the *indegree* of vertex x_i.

Thus, referring to Fig. 1.3, the outdegree of x_6, denoted by $d_0(x_6)$, is $|\Gamma(x_6)| = 2$, and the indegree of x_6 denoted by $d_t(x_6)$, is $|\Gamma^{-1}(x_6)| = 1$.

It is quite obvious that the sum of the outdegrees or indegrees of all the vertices in a graph is equal to the total number of arcs of G i.e.

$$\sum_{i=1}^{n} d_0(x_i) = \sum_{i=1}^{n} d_t(x_i) = m \tag{1.12}$$

where n is the total number of vertices and m the total number of arcs of G.

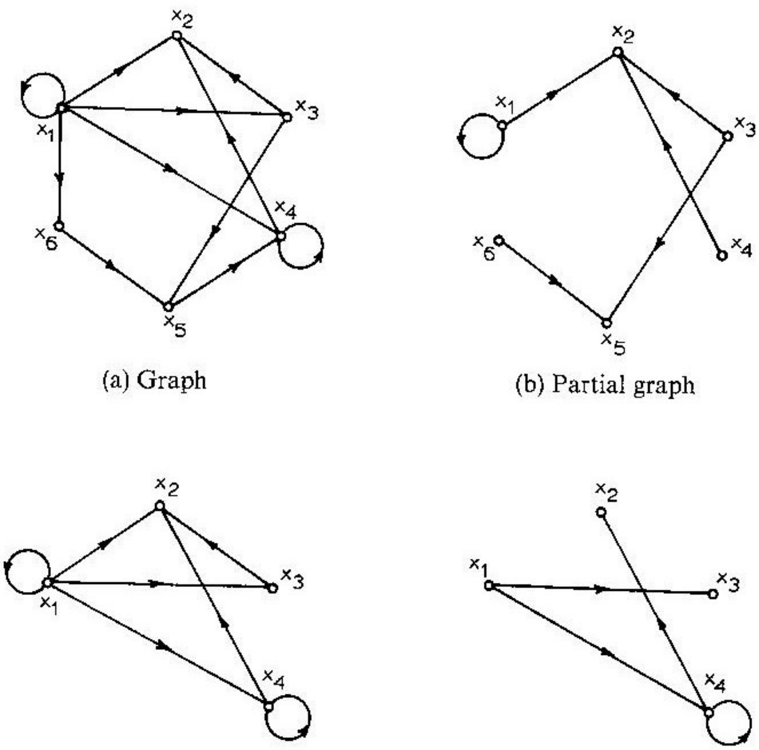
For a nondirected graph $G = (X, \Gamma)$ the degree of a vertex x_i is similarly defined by $d(x_i) \equiv |\Gamma(x_i)|$, and when no confusion can arise it will be written as d_i.

5. Partial Graphs and Subgraphs

Given a graph $G = (X, A)$, a *partial* graph G_p of G is the graph (X, A_p) with $A_p \subset A$. Thus a partial graph is a graph with the same number of vertices but with only a subset of the arcs of the original graph.

If Fig. 1.4(a) represents the graph G, Fig. 1.4(b) is a partial graph G_p.

Given a graph $G = (X, \Gamma)$ a *subgraph* G_s is the graph (X_s, Γ_s) with $X_s \subset X$; and for every $x_i \in X_s$, $\Gamma_s(x_i) = \Gamma(x_i) \cap X_s$. Thus, a subgraph has only a subset X_s of the set of vertices of the original graph but contains all the arcs whose initial and final vertices are both within this subset. It is often very convenient to denote the subgraph G_s simply by $\langle X_s \rangle$ and when no confusion can arise we will use this latter symbolism.



(a) Graph (b) Partial graph

(c) Subgraph (d) Partial subgraph

FIG. 1.4

Fig. 1.4(c) shows a subgraph of the graph in Fig. 1.4(a) containing only vertices x_1, x_2, x_3 and x_4, and those arcs which interconnect them.

The two definitions given above can be combined to define the *partial subgraph*, an example of which is given in Fig. 1.4(d) which shows a partial graph of the subgraph in Fig. 1.4(c).

If a graph represents an entire organization with the vertices representing people and the arcs representing, say, lines of communication, then the graph representing only the more important communication channels of the whole organization is a partial graph; the graph which represents the detailed lines of communication of only a part of the organization (say a division) is a subgraph; and the graph which represents only the important lines of communication within this division is a partial subgraph.

6. Types of Graphs

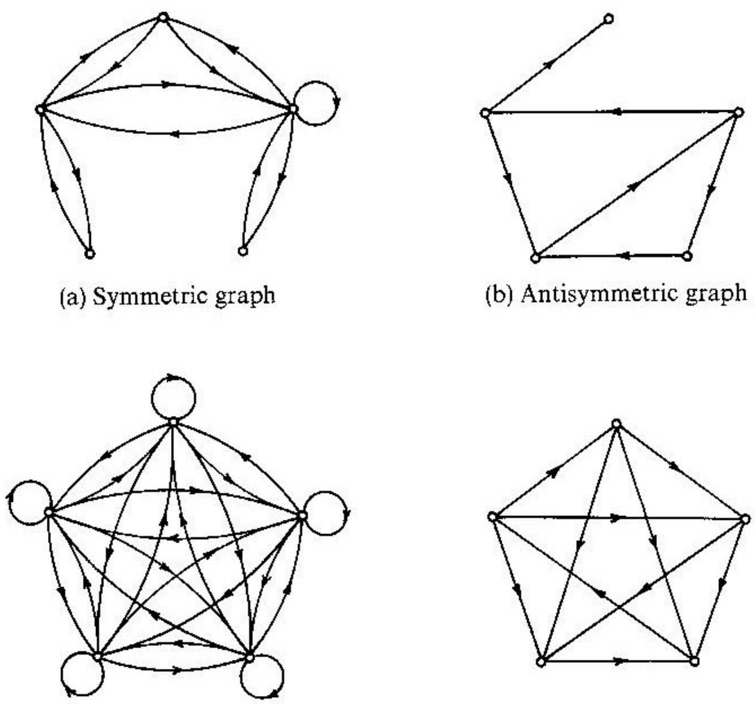
A graph $G = (X, A)$ is said to be *complete* if for every pair of vertices x_i and x_j in X, there exists a link $(\overline{x_i, x_j})$ in $\bar{G} = (X, \bar{A})$ i.e. there must be at least one arc joining every pair of vertices. The complete nondirected graph on n vertices is denoted by K_n.

A graph (X, A) is said to be *symmetric* if, whenever an arc (x_i, x_j) is one of the arcs in the set A of arcs, the opposite arc (x_j, x_i) is also in the set A.

An *antisymmetric* graph is a graph in which whenever an arc $(x_i, x_j) \in A$, the opposite arc $(x_j, x_i) \notin A$. Obviously an antisymmetric graph cannot contain any loops.

Fig. 1.5(a) shows a symmetric and Fig. 1.5(b) an antisymmetric graph.

For example if the vertices of a graph represents a group of people and an arc directed from vertex x_i to vertex x_j means that x_i is the friend or relative of x_j then the graph would be symmetric. On the other hand if an arc directed from x_i to x_j means that x_j is a subordinate of x_i then the resulting graph would be antisymmetric.



(a) Symmetric graph

(b) Antisymmetric graph

(c) Complete symmetric graph

(d) Complete antisymmetric graph (tournament)

FIG. 1.5

Combining the above definitions one can define the *complete symmetric* graph, an example of which is shown in Fig. 1.5(c), and the *complete anti-symmetric* graph, with an example shown in Fig. 1.5(d). This last type of graph is also often referred to as a *tournament*.

A nondirected graph $G = (X, A)$ is said to be *bipartite*, if the set X of its vertices can be partitioned into two subsets X^a and X^b so that all arcs have one terminal vertex in X^a and the other in X^b. A directed graph G is said to be bipartite if its nondirected counterpart $\bar{G}$ is bipartite. It is quite easy to show that:

THEOREM 1. *A nondirected graph G is bipartite if and only if it contains no circuits of odd cardinality.*

Proof. Necessity. Since X is partitioned into X^a and X^b;

$$X^a \cup X^b = X \qquad \text{and} \qquad X^a \cap X^b = \varnothing \tag{1.13}$$

Let an odd cardinality circuit $x_{i_1}, x_{i_2}, \ldots, x_{i_q}, x_{i_1}$ exist and without loss of generality take $x_{i_1} \in X^a$. Since, from the definition, two consecutive vertices on this circuit must belong one to X^a and the other to X^b it follows that $x_{i_2} \in X^b$, $x_{i_3} \in X^a$ etc. and, in general, $x_{i_k} \in X^a$ if k is odd and $x_{i_k} \in X^b$, if k is even. Since we assumed the cardinality of the circuit to be odd, $x_{i_q} \in X^a$ which implies $x_{i_1} \in X^b$. This is a contradiction since $X^a \cap X^b = \varnothing$ and no vertex can belong to both X^a and X^b.

Sufficiency. Assume that no circuits of odd cardinality exist. Choose any vertex x_i, say, and label it "+", the iteratively perform the operations:

Choose any labelled vertex x_i and label all vertices in $\Gamma(x_i)$ with the reverse sign of the label of x_i.

Continue applying this operation until either:

(i) all vertices have been labelled and are consistent, i.e. for any two vertices joined by a link, one is labelled "+" and the other "−", or

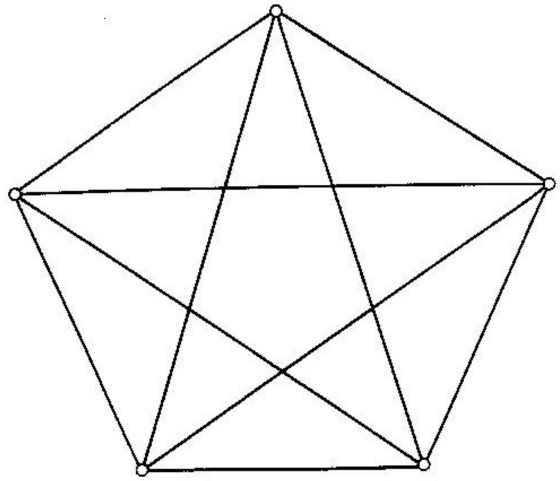
(ii) some vertex, say x_{i_k}, which was labelled with one sign can now be labelled (from a different vertex) with the opposite sign, or

(iii) for all labelled vertices x_i, $\Gamma(x_i)$ is labelled but other unlabelled vertices exist.

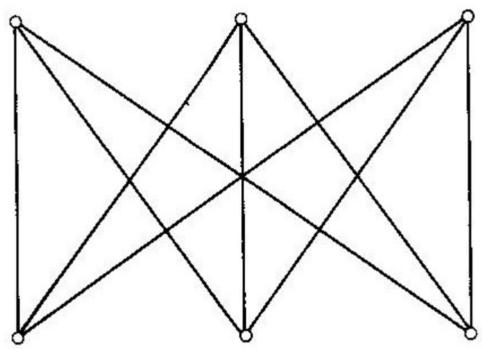
In case (i) let all vertices labelled "+" be in the set X^a and those labelled "−" be in the set X^b. Since all links are between differently labelled vertices, the graph G is bipartite.

In case (ii) vertex x_{i_k} must have received its "+" label along some path (μ_1 say) of vertices, alternatingly labelled "+" and "−", starting from x_{i_1} and finishing at x_{i_k}. Similarly the "−" label on x_{i_k} was obtained along some path μ_2. Let x^* be the last but one (the last being x_{i_k}) vertex common to paths μ_1 and μ_2. If x^* is labelled "+" the part of path μ_2 from x^* to x_{i_k} must be of even, and the part of path μ_2 from x^* to x_{i_k} must be of odd cardinality respectively. The opposite is true if x^* is labelled "−". Hence the circuit consisting of path μ_1 from x^* to x_{i_k} and the reverse part of path μ_2 from x_{i_k} back to x^* is always of odd cardinality. This contradicts the assumption that no odd cardinality circuits exist in G and hence case (ii) is impossible.

Case (iii) implies that there is no link between any labelled and unlabelled



(a)



(b)

FIG. 1.6. The Kuratowski nonplanar graphs. (a) K_5 and (b) $K_{3,3}$

vertex which means that G is disconnected into two or more parts and each part can then be considered in isolation. Thus, only case (i) is eventually possible, and hence the theorem.

When a graph is bipartite and this property needs to be emphasized we will denote the graph as $(X^a \cup X^b, A)$ with equations (1.13) being implied.

A bipartite graph $G = (X^a \cup X^b, A)$ is said to be *complete* if for every two vertices $x_i \in X^a$ and $x_j \in X^b$ there exists a link $(\overline{x_i, x_j})$ in $\bar{G} = (X, \bar{A})$. If $|X^a|$, the number of vertices in set X^a, is r and $|X^b| = s$, then the complete non-directed bipartite graph $G = (X^a \cup X^b, A)$ is denoted by K_{rs}.

A graph $G = (X, A)$ is said to be *planar*, if it can be drawn on a plane (or sphere) in such a way so that no two arcs intersect each other. Figure 1.6(a) shows the complete graph K_5 and Fig. 1.6(b) shows the complete bipartite graph $K_{3,3}$ both of which are known to be *nonplanar* [1, 3]. These two graphs play a central role in planarity and are known as the Kuratowski graphs.

7. Strong Graphs and the Components of a Graph

A graph is said to be strongly connected or *strong* if for any two distinct vertices x_i and x_j, there is at least one path going from x_i to x_j. The above definition implies that any two vertices of a strong graph are mutually reachable.

A graph is said to be unilaterally connected or *unilateral* if for any two distinct vertices x_i and x_j, there is at least one path going from either x_i to x_j, or from x_j to x_i, or both.

A graph is said to be weakly connected or *weak* if there is at least one chain joining every pair of distinct vertices. If for a pair of vertices such a chain does not exist, the graph is said to be *disconnected* [2].

Considering, for example, the graph shown in Fig. 1.7(a) it is easy to check that this is strong. The graph shown in Fig. 1.7(b) on the other hand is not strong (there is no path going from x_1 to x_3), but is unilateral. The graph shown in Fig. 1.7(c) is neither strong nor unilateral—since there is no path going from x_2 to x_5 nor one from x_5 to x_2. It is, however, weak. Finally, the graph of Fig. 1.7(d) is disconnected.

Given any property P to characterize a graph, a *maximal subgraph* $\langle \hat{X}_s \rangle$ if a graph G with respect to that property, is a subgraph which has this property, and there is no other subgraph $\langle X_s \rangle$ with $X_s \supset \hat{X}_s$ and which also has the property. Thus, if the property P is strong-connectedness, then a maximal strong subgraph of G is a strong subgraph of G which is not contained in any other strong subgraph. Such a subgraph is called a *strong component* of G. Similarly a *unilateral component* is a maximal unilateral subgraph and a *weak component* is a maximal weak subgraph.

For example, in the graph G shown in Fig. 1.7(b), the subgraph $\langle \{x_1, x_4, x_5, x_6\} \rangle$ is a strong component of G. On the other hand $\langle \{x_1, x_6\} \rangle$ and $\langle \{x_1, x_5, x_6\} \rangle$ are not strong components—even though they are strong subgraphs—since these are contained in $\langle \{x_1, x_4, x_5, x_6\} \rangle$ and are, therefore, not maximal. In the graph shown in Fig. 1.7(c) the subgraph $\langle \{x_1, x_4, x_5\} \rangle$ is a unilateral component. In the graph shown in Fig. 1.7(d) the subgraphs $\langle \{x_1, x_5, x_6\} \rangle$ and $\langle \{x_2, x_3, x_4\} \rangle$ are both weak components and are the only two such components.

It is quite obvious from the definitions that unilateral components are not necessarily pairwise vertex-disjoint. A strong component must be contained in at least one unilateral component and a unilateral component contained in a weak component for any given graph G.

8. Matrix Representations

A convenient way of representing a graph algebraically is by the use of matrices, as follows.

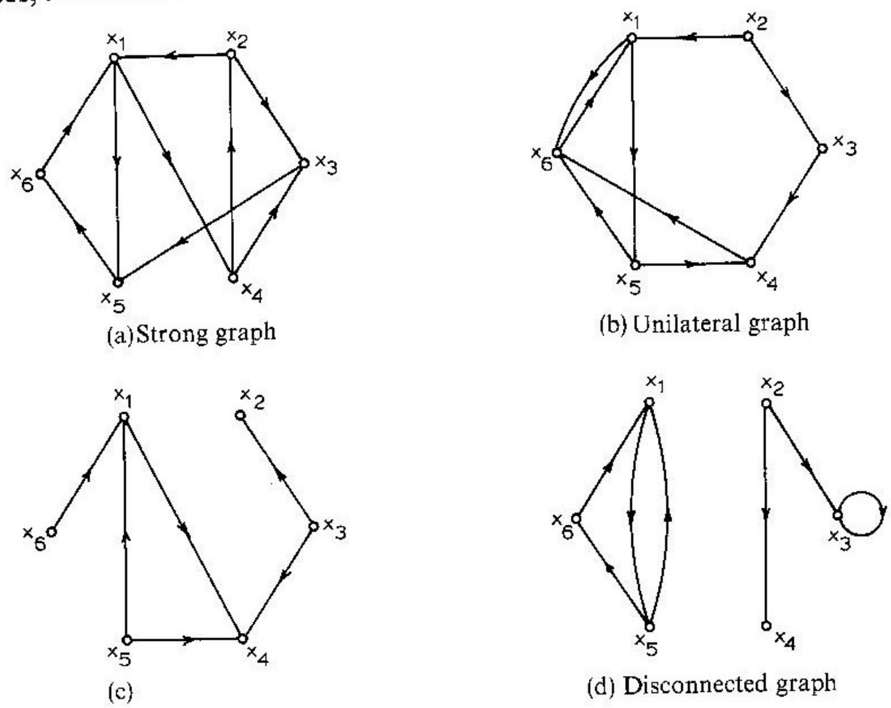


FIG. 1.7.

8.1 The adjacency matrix

Given a graph G, its *adjacency* matrix is denoted by $\mathbf{A} = [a_{ij}]$ and is given by:

$$a_{ij} = 1 \text{ if arc } (x_i, x_j) \text{ exists in } G$$

$$a_{ij} = 0 \text{ if arc } (x_i, x_j) \text{ does not exist in } G.$$

Thus, the adjacency matrix of the graph shown in Fig. 1.8 is:

$$\mathbf{A} = \begin{array}{c} \\ x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \end{array} \begin{array}{c} \begin{array}{cccccc} x_1 & x_2 & x_3 & x_4 & x_5 & x_6 \end{array} \\ \left[\begin{array}{cccccc} 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 1 \end{array}\right] \end{array}$$

The adjacency matrix defines completely the structure of the graph. For example, the sum of all the elements in row x_i of the matrix gives the out-degree of vertex x_i and the sum of the elements in column x_i gives the in-degree of vertex x_i. The set of columns which have an entry of 1 in row x_i is the set $\Gamma(x_i)$, and the set of rows which have an entry of 1 in column x_i is the set $\Gamma^{-1}(x_i)$.

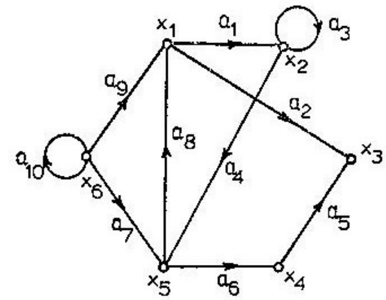


FIG. 1.8.

Consider the adjacency matrix raised to the second power. An element, $a_{ik}^{(2)}$ say, of matrix $\mathbf{A}^2$ is given by:

$$a_{ik}^{(2)} = \sum_{j=1}^{n} a_{ij}a_{jk} \tag{1.14}$$

Each term in the summation of equation (1.14) has the value 1 only if a_{ij} and a_{jk} are both 1, otherwise it has the value 0. Since $a_{ij} = a_{jk} = 1$ implies a path of cardinality 2 from vertex x_i to vertex x_k via vertex x_j, the term $a_{ik}^{(2)}$ is then the total number of cardinality 2 paths from x_i to x_k.

Similarly, if $a_{ik}^{(p)}$ is an element of $\mathbf{A}^p$, then $a_{ik}^{(p)}$ is the number of paths (not necessarily simple or elementary) of cardinality p from x_i to x_k.

8.2 The incidence matrix

Given a graph G of n vertices and m arcs, the *incidence* matrix of G is denoted by $\mathbf{B} = [b_{ij}]$ and is an $n \times m$ matrix defined as follows.

$$b_{ij} = 1 \text{ if } x_i \text{ is the initial vertex of arc } a_j$$

$$b_{ij} = -1 \text{ if } x_i \text{ is the final vertex of arc } a_j$$

and $b_{ij} = 0$ if x_i is not a terminal vertex of arc a_j or if a_j is a loop.

For the graph shown in Fig. 1.8, the incidence matrix is:

$$\mathbf{B} = \begin{array}{c|cccccccccc|} & a_1 & a_2 & a_3 & a_4 & a_5 & a_6 & a_7 & a_8 & a_9 & a_{10} \\ \hline x_1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & -1 & -1 & 0 \\ x_2 & -1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ x_3 & 0 & -1 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 \\ x_4 & 0 & 0 & 0 & 0 & 1 & -1 & 0 & 0 & 0 & 0 \\ x_5 & 0 & 0 & 0 & -1 & 0 & 1 & -1 & 1 & 0 & 0 \\ x_6 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 \\ \hline \end{array}$$

Since each arc is adjacent to exactly two vertices, each column of the incidence matrix contains one 1 and one -1 entry, except when the arc forms a loop in which case it contains only zero entries.

If G is a nondirected graph, then the incidence matrix is defined as above except that all entries of -1 are now changed to $+1$.

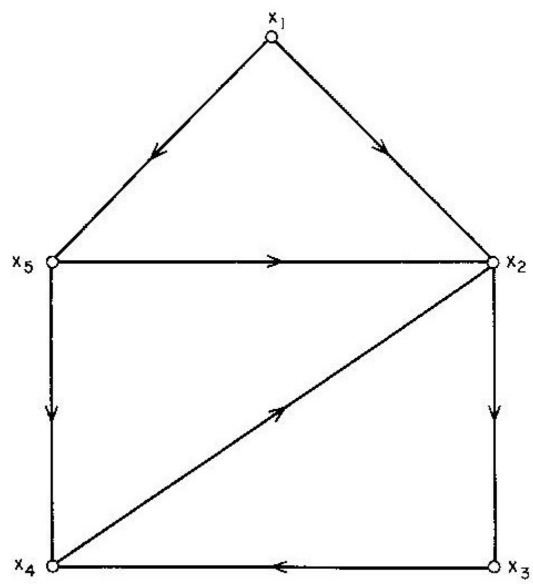


FIG. 1.9.

9. Problems P1

1. For the graph of Fig. 1.9 find:
   - (a) $\Gamma(x_2)$
   - (b) $\Gamma^{-1}(x_2)$
   - (c) $\Gamma^2(x_2)$
   - (d) $\Gamma^{-2}(x_2)$
   - (e) $d_0(x_2)$
   - (f) $d_t(x_2)$
   - (g) the adjacency matrix **A**
   - (h) the incidence matrix **B**

2. For the graph $G = (X, A)$ of Fig 1.9 draw:
   (a) The subgraph $\langle\{x_1, x_2, x_4, x_5\}\rangle$
   (b) The partial graph (X, A'), where $(x_i, x_j) \in A'$ if and only if $i + j$ is odd.
   (c) The partial graph as defined in (b) of the subgraph in (a).
3. For a nondirected graph prove that the number of vertices of odd degree is even. (Zero is an even number).
4. Show that any complete symmetric graph contains a Hamiltonian circuit.
5. Show that the rank of the incidence matrix **B** of a connected graph is $n - 1$, and hence prove that the rank of **B** for a graph with P (weak) components is $n - P$.
6. Prove that a nondirected connected graph remains connected after the removal of a link if and only if that link is part of some circuit.
7. Prove that a nondirected connected graph with n vertices
   (a) Contains at least $n - 1$ links
   (b) If it contains more than $n - 1$ links then it has at least one circuit.

10. References

1. Berge, C., (1962). "The theory of graphs", Methuen, London.
2. Harary, F., Norman, R. Z. and Cartwright, D. (1965). "Structural models: An introduction to the theory of directed graphs," Wiley, New York.
3. Kuratowski, G. (1930). Sur le problème des courbes gauches en topologie, *Fund. Match.*, 15–16, p. 271.
4. Roy, B. (1969, 1970). Algebre moderne et theorie des graphes, Vol. 1 (1969), Vol. 2 (1970), Dunod, Paris.