Directed Graph

Prepared by Maksuda Rabeya

9.2 DIRECTED GRAPHS

A *directed graph* G or *digraph* (or simply *graph*) consists of two things:

(i) A set V whose elements are called *vertices*, *nodes*, or *points*.

(ii) A set E of *ordered* pairs (u, v) of vertices called *arcs* or *directed edges* or simply *edges*.

We will write $G(V, E)$ when we want to emphasize the two parts of G. We will also write $V(G)$ and $E(G)$ to denote, respectively, the set of vertices and the set of edges of a graph G. (If it is not explicitly stated, the context usually determines whether or not a graph G is a directed graph.)

Suppose $e = (u, v)$ is a directed edge in a digraph G. Then the following terminology is used:

(a) e *begins* at u and *ends* at v.
(b) u is the *origin* or *initial point* of e, and v is the *destination* or *terminal point* of e.
(c) v is a *successor* of u.
(d) u is *adjacent to* v, and v is *adjacent from* u.

If $u = v$, then e is called a *loop*.

The set of all successors of a vertex u is important; it is denoted and formally defined by

$$\mathrm{succ}(u) = \{v \in V \mid \text{there exists an edge } (u, v) \in E\}$$

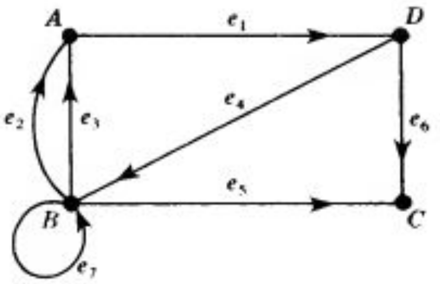
It is called the *successor list* or *adjacency list* of u.

EXAMPLE 9.1

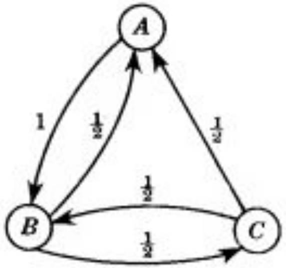
(a) Consider the directed graph G pictured in Fig. 9-1(a). It consists of four vertices, A, B, C, D, that is, $V(G) = \{A, B, C, D\}$ and the seven following edges:

$$E(G) = \{e_1, e_2, \ldots, e_7\} = \{(A, D), (B, A), (B, A), (D, B), (B, C), (D, C), (B, B)\}$$

The edges e_2 and e_3 are said to be *parallel* since they both begin at B and end at A. The edge e_7 is a *loop* since it begins and ends at B.



(a)



(b)

Fig. 9-1

(b) Suppose three boys, A, B, C, are throwing a ball to each other such that A always throws the ball to B, but B and C are just as likely to throw the ball to A as they are to each other. This dynamic system is pictured in Fig. 9-1(b) where edges are labeled with the respective probabilities, that is, A throws the ball to B with probability 1, B throws the ball to A and C each with probability 1/2, and C throws the ball to A and B each with probability 1/2.

Subgraphs

Let $G = G(V, E)$ be a directed graph, and let V' be a subset of the set V of vertices of G. Suppose E' is a subset of E such that the endpoints of the edges in E' belong to V'. Then $H(V', E')$ is a directed graph, and it is called a *subgraph* of G. In particular, if E' contains all the edges in E whose endpoints belong to V', then $H(V', E')$ is called the subgraph of G *generated* or *determined* by V'. For example, for the graph $G = G(V, E)$ in Fig. 9-1(a), $H(V', E')$ is the subgraph of G determine by the vertex set V' where

$$V' = \{B, C, D\} \quad \text{and} \quad E' = \{e_4, e_5, e_6, e_7\} = \{(D, B), (B, C), (D, C), (B, B)\}$$

Degrees

Suppose G is a directed graph. The *outdegree* of a vertex v of G, written outdeg(v), is the number of edges beginning at v, and the *indegree* of v, written indeg(v), is the number of edges ending at v. Since each edge begins and ends at a vertex we immediately obtain the following theorem.

Theorem 9.1: The sum of the outdegrees of the vertices of a digraph G equals the sum of the indegrees of the vertices, which equals the number of edges in G.

A vertex v with zero indegree is called a *source*, and a vertex v with zero outdegree is called a *sink*.

EXAMPLE 9.2 Consider the graph G in Fig. 9-1(a). We have:

$$\begin{array}{cccc} \text{outdeg}\,(A) = 1, & \text{outdeg}\,(B) = 4, & \text{outdeg}\,(C) = 0, & \text{outdeg}\,(D) = 2, \\ \text{indeg}\,(A) = 2, & \text{indeg}\,(B) = 2, & \text{indeg}\,(C) = 2, & \text{indeg}\,(D) = 1. \end{array}$$

As expected, the sum of the outdegrees equals the sum of the indegrees, which equals the number 7 of edges. The vertex C is a sink since no edge begins at C. The graph has no sources.

Paths

Let G be a directed graph. The concepts of path, simple path, trail, and cycle carry over from nondirected graphs to the directed graph G except that the directions of the edges must agree with the direction of the path. Specifically:

(i) A *(directed) path* P in G is an alternating sequence of vertices and directed edges, say,

$$P = (v_0, e_1, v_1, e_2, v_2, \ldots, e_n, v_n)$$

such that each edge e_i begins at v_{i-1} and ends at v_i. If there is no ambiguity, we denote P by its sequence of vertices or its sequence of edges.

(ii) The *length* of the path P is n, its number of edges.

(iii) A *simple path* is a path with distinct vertices. A *trail* is a path with distinct edges.

(iv) A *closed path* has the same first and last vertices.

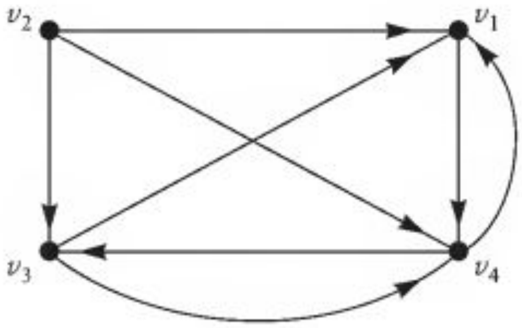
(v) A *spanning path* contains all the vertices of G.

(vi) A *cycle* (or *circuit*) is a closed path with distinct vertices (except the first and last).

(vii) A *semipath* is the same as a path except the edge e_i may begin at v_{i-1} or v_i and end at the other vertex. *Semitrails* and *semisimple paths* are analogously defined.

A vertex v is *reachable* from a vertex u if there is a path from u to v. If v is reachable from u, then (by eliminating redundant edges) there must be a simple path from u to v.

EXAMPLE 9.6 Let G be the directed graph in Fig. 9-4(a) with vertices v_1, v_2, v_3, v_4. Then the adjacency matrix A of G appears in Fig. 9-4(b). Note that the number of 1's in A is equal to the number (eight) of edges.



(a)

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

(b)

Fig. 9-4

Consider the powers $A, A^2, A^3, \ldots$ of the adjacency matrix $A = [a_{ij}]$ of a graph G. Let

$$a_K(i, j) = \text{the } ij \text{ entry in the matrix } A^K$$

Note that $a_1(i, j) = a_{ij}$ gives the number of paths of length 1 from vertex v_i to vertex v_j. One can show that $a_2(i, j)$ gives the number of paths of length 2 from v_i to v_j. In fact, we prove in Problem 9.17 the following general result.

EXAMPLE 9.7 Consider again the graph G and its adjacency matrix A appearing in Fig. 9-4. The powers A^2, A^3, and A^4 of A follow:

$$A^2 = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 2 & 0 & 1 & 2 \\ 1 & 0 & 1 & 1 \\ 1 & 0 & 0 & 2 \end{bmatrix}, \quad A^3 = \begin{bmatrix} 1 & 0 & 0 & 2 \\ 3 & 0 & 2 & 3 \\ 2 & 0 & 1 & 2 \\ 2 & 0 & 2 & 1 \end{bmatrix}, \quad A^4 = \begin{bmatrix} 2 & 0 & 2 & 1 \\ 5 & 0 & 3 & 5 \\ 3 & 0 & 2 & 3 \\ 3 & 0 & 1 & 4 \end{bmatrix}$$

Observe that $a_2(4, 1) = 1$, so there is a path of length 2 from v_4 to v_1. Also, $a_3(2, 3) = 2$, so there are two paths of length 3 from v_2 to v_3; and $a_4(2, 4) = 5$, so there are five paths of length 4 from v_2 to v_4.

Remark: Let A be the adjacency matrix of a graph G, and let B_r be the matrix defined by:

$$B_r = A + A^2 + A^3 + \cdots + A^r$$

Then the ij entry of the matrix B_r gives the number of paths of length r or less from vertex v_i to vertex v_j.

Path Matrix

Let $G = G(V, E)$ be a simple directed graph with m vertices $v_1, v_2, \ldots, v_m$. The *path matrix* or *reachability matrix* of G is the m-square matrix $P = [p_{ij}]$ defined as follows:

$$p_{ij} = \begin{cases} 1 & \text{if there is a path from to } v_i \text{ to } v_j \\ 0 & \text{otherwise} \end{cases}$$

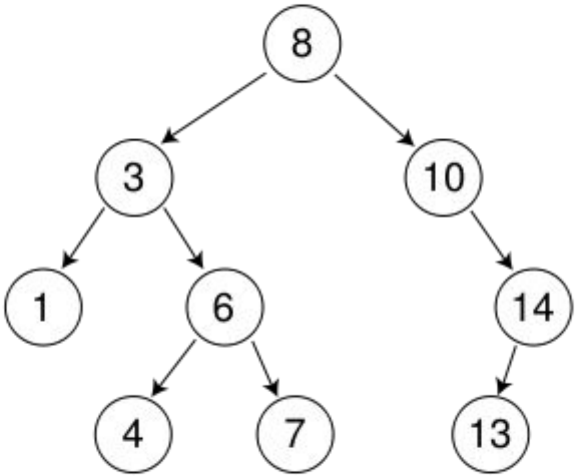
(The path matrix P may be viewed as the transitive closure of the relation E on V.)

EXAMPLE 9.8 Consider the graph G and its adjacency matrix A appearing in Fig. 9-4. Here G has $m = 4$ vertices. Adding the matrix A and matrices A^2, A^3, A^4 in Example 9.7, we obtain the following matrix B_4 and also path (reachability) matrix P by replacing the nonzero entries in B_4 by 1:

$$B_4 = \begin{bmatrix} 4 & 0 & 3 & 4 \\ 11 & 0 & 7 & 11 \\ 7 & 0 & 4 & 7 \\ 7 & 0 & 4 & 7 \end{bmatrix} \quad \text{and} \quad P = \begin{bmatrix} 1 & 0 & 1 & 1 \\ 1 & 0 & 1 & 1 \\ 1 & 0 & 1 & 1 \\ 1 & 0 & 1 & 1 \end{bmatrix}$$

Examining the matrix B_4 or P, we see zero entries; hence G is not strongly connected. In particular, we see that the vertex v_2 is not reachable from any of the other vertices.

Breadth first search



Depth first search