

# 第4章 图结构及其应用算法





## 图论——欧拉

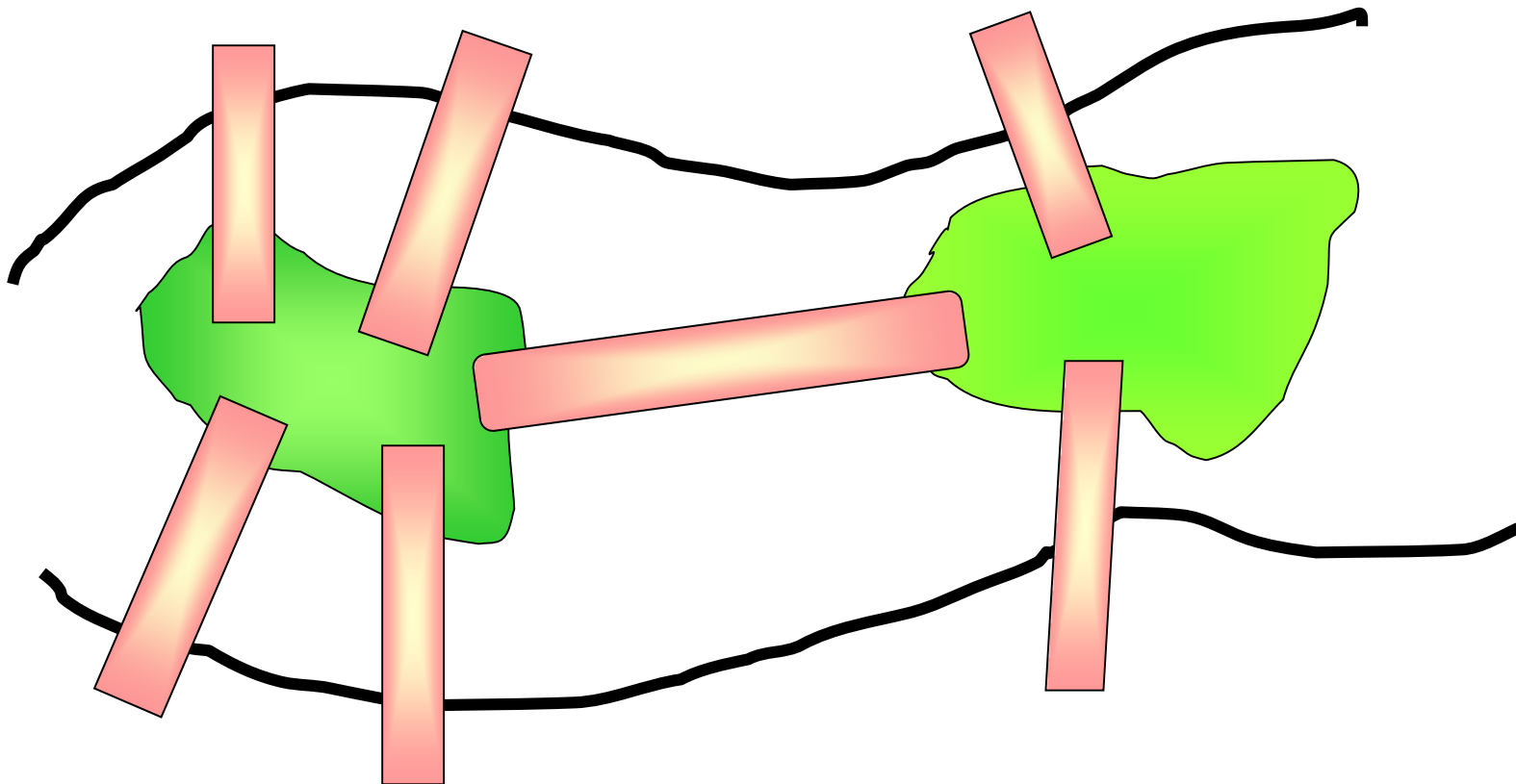


欧拉1707年出生在瑞士的巴塞尔城，19岁开始发表论文，直到76岁。几乎每一个数学领域都可以看到欧拉的名字，从初等几何的欧拉线，多面体的欧拉定理，立体解析几何的欧拉变换公式，四次方程的欧拉解法到数论中的欧拉函数，微分方程的欧拉方程，级数论的欧拉常数，变分学的欧拉方程，复变函数的欧拉公式等等。据统计他那不倦的一生，共写下了886本书籍和论文，其中分析、代数、数论占40%，几何占18%，物理和力学占28%，天文学占11%，弹道学、航海学、建筑学等占3%。1733年，年仅26岁的欧拉担任了彼得堡科学院数学教授。1741年到柏林担任科学院物理数学所所长，直到1766年，重回彼得堡，没有多久，完全失明。欧拉在数学上的建树很多，对著名的哥尼斯堡七桥问题的解答开创了图论的研究。





# 哥尼斯堡七桥问题



➡ 从某个陆地区域出发，是否存在一条能够经过每座桥一次且仅一次，最后回到出发地？





# 学习目标

- 图结构是一种非线性结构，反映了数据对象之间的**任意关系**，在计算机科学、数学和工程中有着非常广泛的应用。
- 了解图的**定义及相关的术语**，掌握图的**逻辑结构及其特点**；
- 了解图的**存储方法**，重点掌握图的**邻接矩阵**和**邻接表**存储结构；
- 掌握图的**遍历方法**，重点掌握**图的遍历算法的实现**；
- 了解**图的应用**，重点掌握**最小生成树算法**、**最短路径算法**、**拓扑排序**和**关键路径算法**的**基本思想**、**算法原理**和**实现过程**。





# 主要内容

- 4.1 图的基本概念
- 4.2 图的存储结构
- 4.3 图的遍历
- 4.4 图与树的关系、最小生成树算法
- 4.5 无向图的双连通性
- 4.6 拓扑排序算法
- 4.7 关键路径算法
- 4.8 最短路径算法
- 本章小结





## 4.1 基本定义

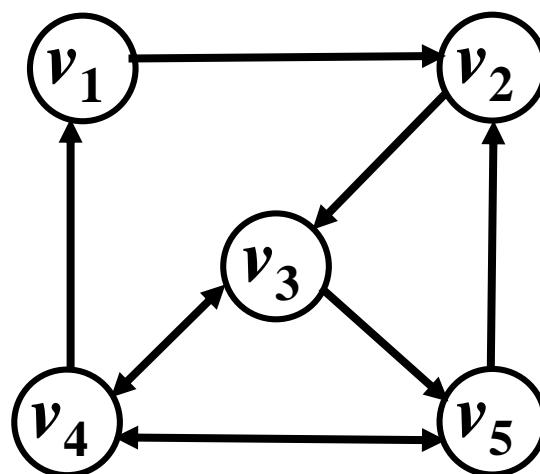
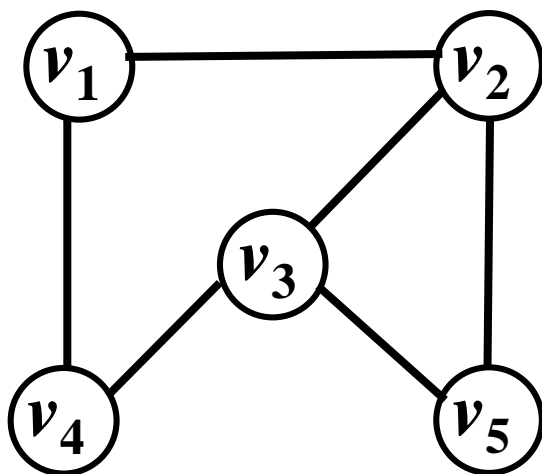
### 定义1 图(Graph)

➡ 图是由**顶点 (vertex)** 的**有穷非空**集合和顶点之间**边 (edge)** 的集合组成的一种数据结构，通常表示为：

$$G = (V, E)$$

其中：**G**表示一个图，**V**是图**G**中顶点的集合，**E**是图**G**中顶点之间边的集合。

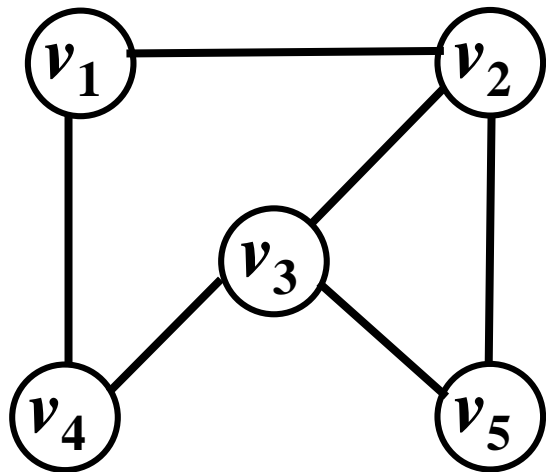
**顶点**表示**数据对象**；**边**表示**数据对象之间的关系**。





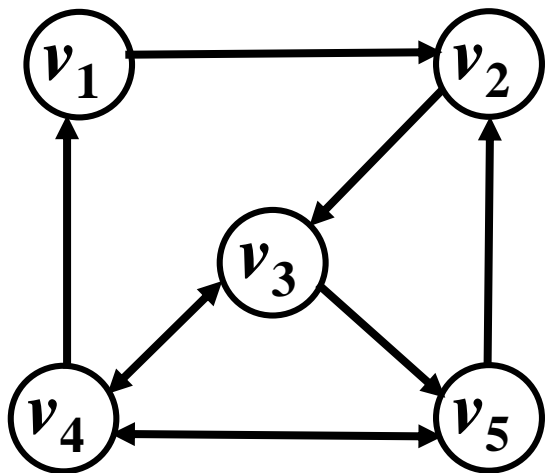
## 4.1 基本定义(cont.)

### 定义1 图



#### 无向图:

- 若顶点 $v_i$ 和 $v_j$ 之间的边没有方向，则称这条边为**无向边**，表示为 $(v_i, v_j)$ 。
- 如果图的任意两个顶点之间的边都是无向边，则称该图为**无向图**。



#### 有向图:

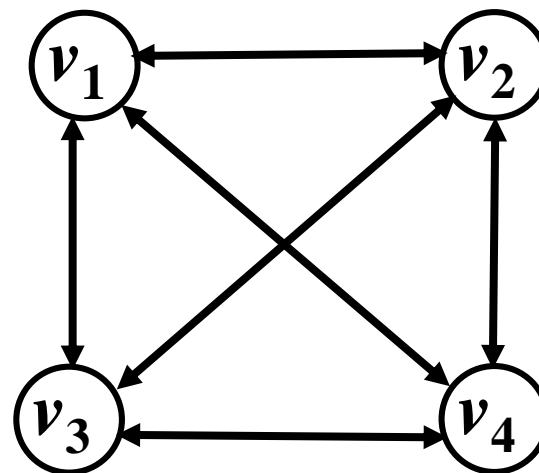
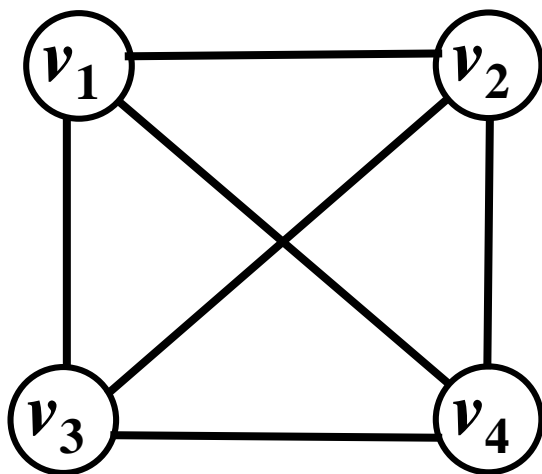
- 若顶点 $v_i$ 和 $v_j$ 之间的边都有方向，则称这条边为**有向边(弧)**，表示为 $\langle v_i, v_j \rangle$ 。
- 如果图的任意两个顶点之间的边都是有向边，则称该图为**有向图**。





## 4.1 基本定义(cont.)

- **无向完全图**：在无向图中，如果任意两个顶点之间都存在边，则称该图为无向完全图。
- **有向完全图**：在有向图中，如果任意两个顶点之间都存在方向相反的两条弧，则称该图为有向完全图。



- 含有 $n$ 个顶点的无向完全图有多少条边？
- 含有 $n$ 个顶点的有向完全图有多少条弧？



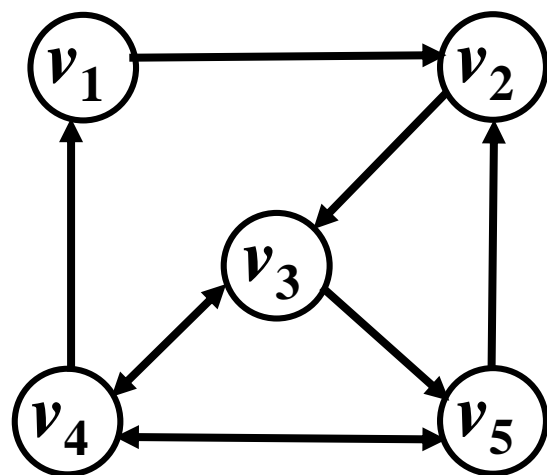
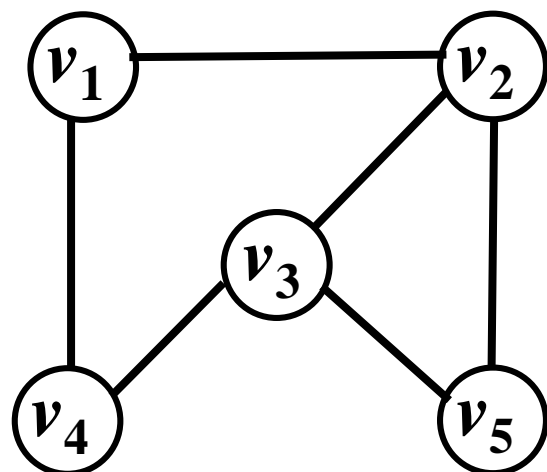




## 4.1 基本定义(cont.)

### 定义1 图

#### ➡ 邻接、依附



■ 在**无向图**中，对于任意两个顶点 $v_i$ 和顶点 $v_j$ ，若存在边 $(v_i, v_j)$ ，则称顶点 $v_i$ 和顶点 $v_j$ **相邻**，互为**邻接点**，同时称边 $(v_i, v_j)$ **依附于**顶点 $v_i$ 和顶点 $v_j$ 。

■ 如： $v_2$ 的邻接点： $v_1, v_3, v_5$

■ 在**有向图**中，对于任意两个顶点 $v_i$ 和顶点 $v_j$ ，若存在有向边 $\langle v_i, v_j \rangle$ ，则称顶点 $v_i$ **邻接到**顶点 $v_j$ ，顶点 $v_j$ **邻接于**顶点 $v_i$ ，同时称弧 $\langle v_i, v_j \rangle$ **依附于**顶点 $v_i$ 和顶点 $v_j$ 。

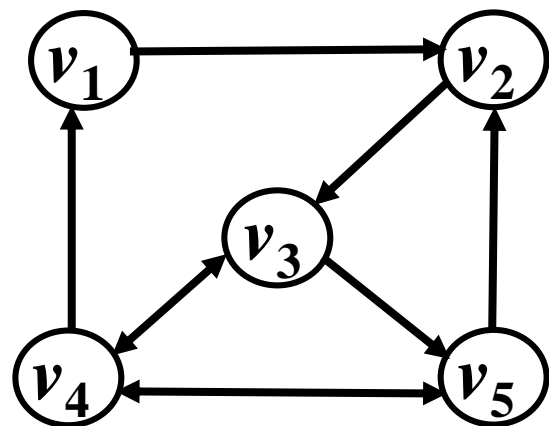
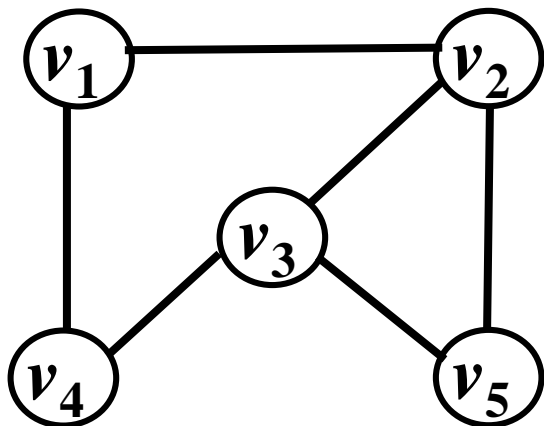
■ 如： $v_1$ 邻接到 $v_2$ ， $v_1$ 邻接于 $v_4$





## 4.1 基本定义(cont.)

### 定义2 度(Dgree)



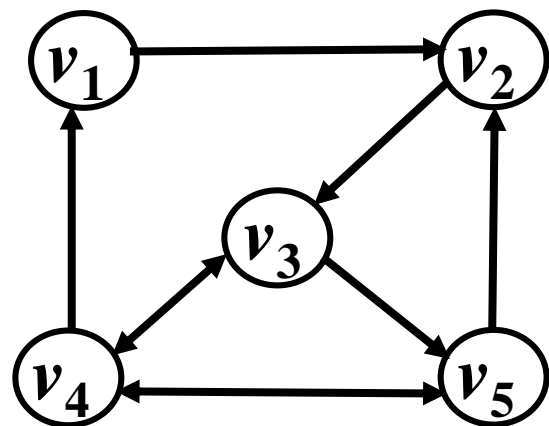
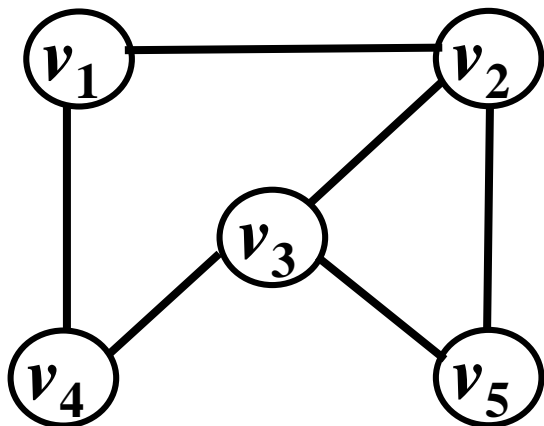
- **顶点的度**: 在无向图中, 顶点 $v$ 的**度**是指依附于该顶点的边数, 通常记为 $D(v)$ 。
- **顶点的入度**: 在有向图中, 顶点 $v$ 的**入度**是指以该顶点为弧头的弧的数目, 记为 $ID(v)$ ;
- **顶点的出度**: 在有向图中, 顶点 $v$ 的**出度**是指以该顶点为弧尾的弧的数目, 记为 $OD(v)$ 。
- 在有向图中,  $D(v) = ID(v) + OD(v)$





## 4.1 基本定义(cont.)

### 定义2 度(Dgree)



- 在具有 $n$ 个顶点、 $e$ 条边的无向图 $G$ 中，各顶点的度之和与边数之和的关系？

$$\sum_{i=1}^n D(v_i) = 2e$$

- 在具有 $n$ 个顶点、 $e$ 条边的有向图 $G$ 中，各顶点的入度之和与各顶点的出度之和的关系？与边数之和的关系？

$$\sum_{i=1}^n ID(v_i) = \sum_{i=1}^n OD(v_i) = e$$





## 4.1 基本定义(cont.)

### 定义3 路径 (Path) 和路径长度、简单路和简单回路

- 在**无向图**  $G=(V, E)$  中, 若存在一个**顶点序列**  $v_p, v_{i1}, v_{i2}, \dots, v_{im}, v_q$ , 使得  $(v_p, v_{i1}), (v_{i1}, v_{i2}), \dots, (v_{im}, v_q) \in E(G)$ , 则称顶点  $v_p$  路到  $v_q$  有一条**路径**。
- 在有向图  $G=(V, E)$  中, 若存在一个**顶点序列**  $v_p, v_{i1}, v_{i2}, \dots, v_{im}, v_q$ , 使得有向边  $\langle v_p, v_{i1} \rangle, \langle v_{i1}, v_{i2} \rangle, \dots, \langle v_{im}, v_q \rangle \in E(G)$ , 则称顶点  $v_p$  路到  $v_q$  有一条**有向路径**。
- **非带权图**的**路径长度**是指此路径上边的条数。
- **带权图**的**路径长度**是指路径上各边的权之和。
- **简单路径**: 若路径上各顶点  $v_1, v_2, \dots, v_m$  均互不相同(**第一个顶点和最后一个顶点可以相同**), 则称这样的路径为简单路径。
- **简单回路**: 若路径上第一个顶点  $v_1$  与最后一个顶点  $v_m$  重合, 则称这样的简单路径为**简单回路或环**。
- 一条回路的长度至少为1 (无向图为3), 且起点和终点相同的简单路径。



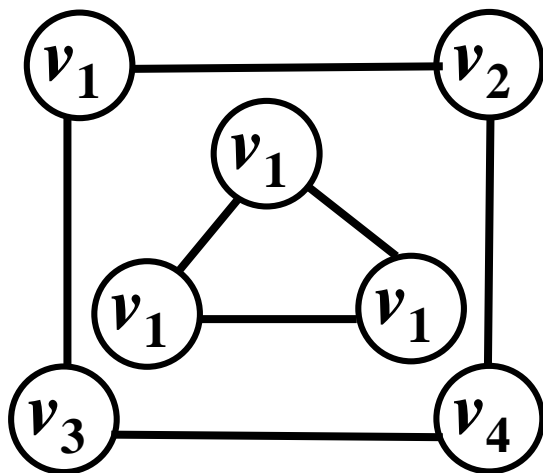


## 4.1 基本定义(cont.)

### 定义4 图的连通性

#### ➡ 连通图与连通分量

- **顶点的连通性**: 在无向图中, 若从顶点 $v_i$ 到顶点 $v_j$  ( $i \neq j$ ) 有路径, 则称顶点 $v_i$ 与 $v_j$ 是**连通的**。
- **连通图**: 如果一个无向图中任意一对顶点都是连通的, 则称此图是**连通图**。
- **连通分量**: 非连通图的极大连通子图叫做**连通分量**。



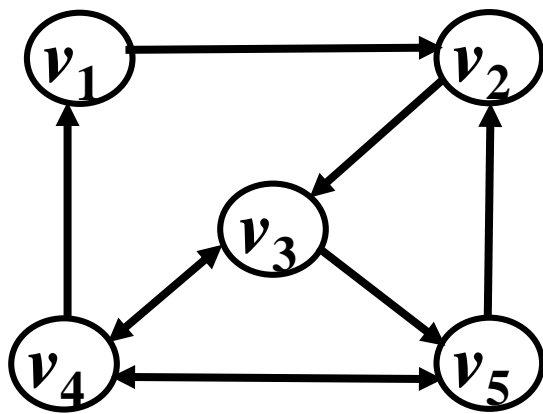


## 4.1 基本定义(cont.)

### 定义4 图的连通性

#### ➤ 强连通图与强连通分量

- **顶点的强连通性**: 在有向图中, 若对于每一对顶点 $v_i$ 和 $v_j$  ( $i \neq j$ ), 都存在一条从 $v_i$ 到 $v_j$ 和从 $v_j$ 到 $v_i$ 的**有向**路径, 则称顶点 $v_i$ 与 $v_j$ 是**强连通的**。
- **强连通图**: 如果一个有向图中任意一对顶点都是强连通的, 则称此有向图是**强连通图**。
- **强连通分量**: 非强连通图的极大强连通子图叫做**强连通分量**





## 4.1 基本定义(cont.)

### 图的操作

设图 $G=(V,E)$ ，图上定义的基本操作如下：

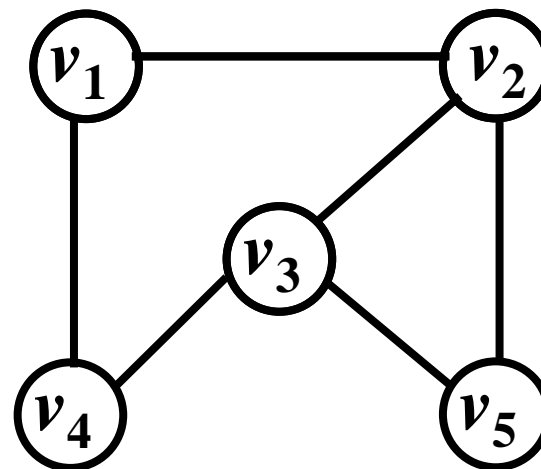
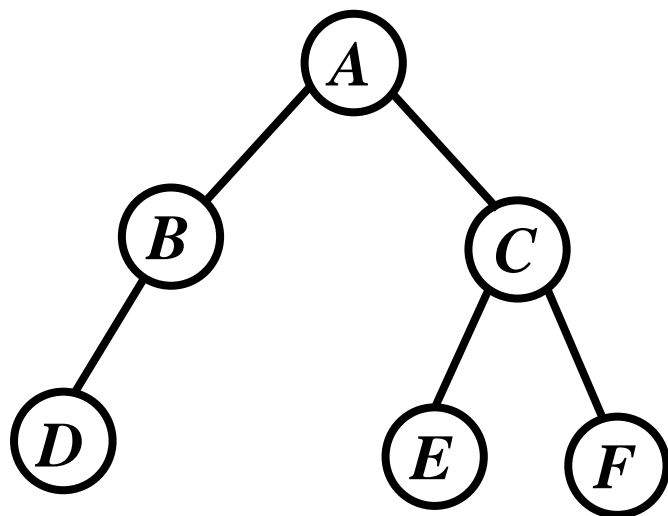
- **NewNode ( G )**: 建立一个新顶点,  $V=V \cup \{v\}$
- **DelNone ( G, v )**: 删除顶点 $v$ 以及与之相关联的所有边
- **SetSucc ( G, v1, v2 )**: 增加一条边,  $E = E \cup (v1,v2)$ ,  $V=V$
- **DelSucc ( G, v1, v2 )**: 删除边  $(v1,v2)$ ,  $V$ 不变
- **Succ ( G, v1, v2 )**: 求出 $v$ 的所有直接后继结点
- **Pred ( G, v )**: 求出 $v$ 的所有直接前导结点
- **IsEdge ( G, v1, v2 )**: 判断  $(v1,v2) \in E$
- **FirstAdjVex( G , v )**: 顶点 $v$  的第一个邻接顶点
- **NextAdjVex( G, v, w )**: 顶点 $v$  的某个邻接点 $w$ 的下一个邻接顶点。





## 4.1 基本定义(cont.)

### 不同逻辑结构之间的比较



- 在线性结构中，数据元素之间仅具有**线性关系(1:1)**；
- 在树型结构中，结点之间具有**层次关系(1:m)**；
- 在图型结构中，任意两个顶点之间**都可能有关系(m:n)**。

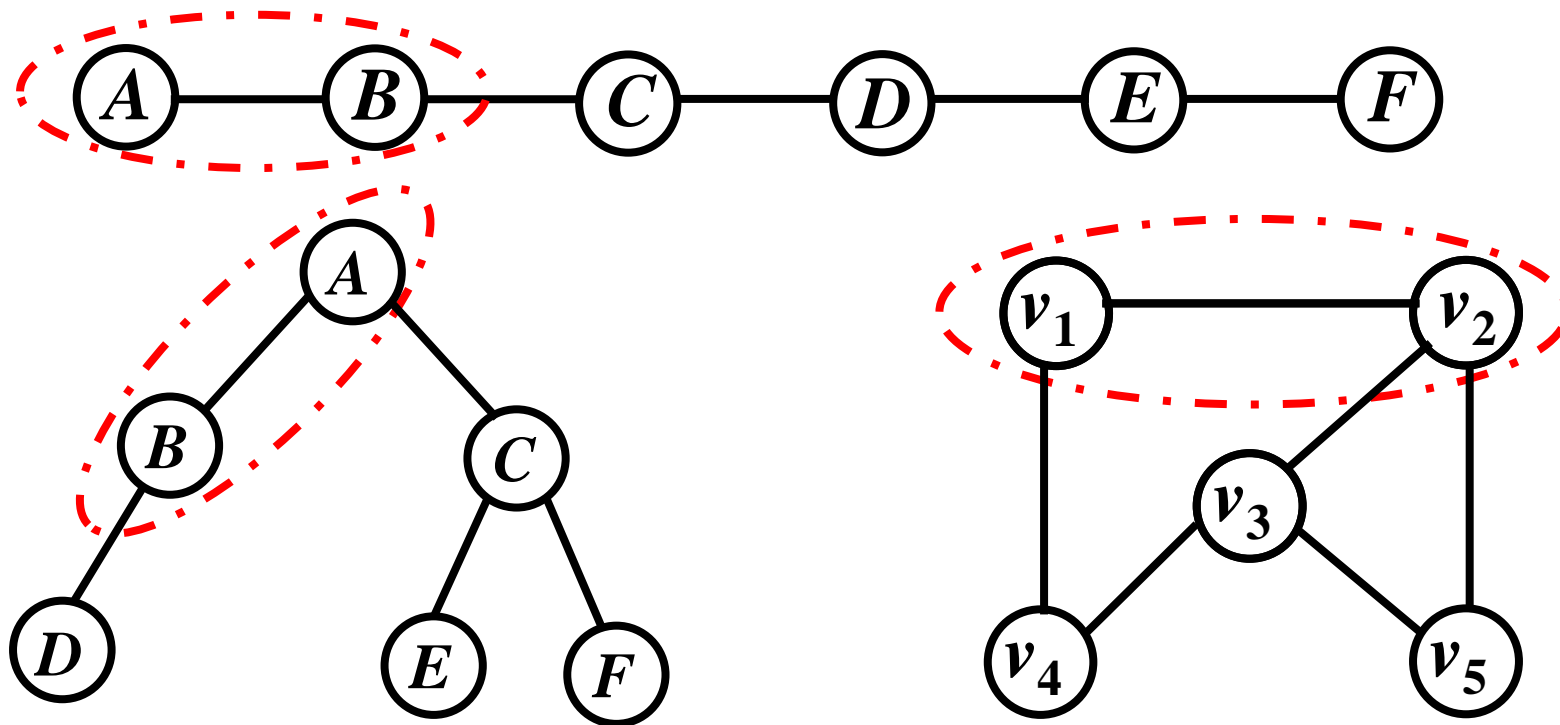






## 4.1 基本定义(cont.)

不同逻辑结构之间的比较



- 在线性结构中，元素之间的关系为**前驱**和**后继**；
- 在树型结构中，结点之间的关系为**双亲**和**孩子**；
- 在图型结构中，顶点之间的关系为**邻接**。





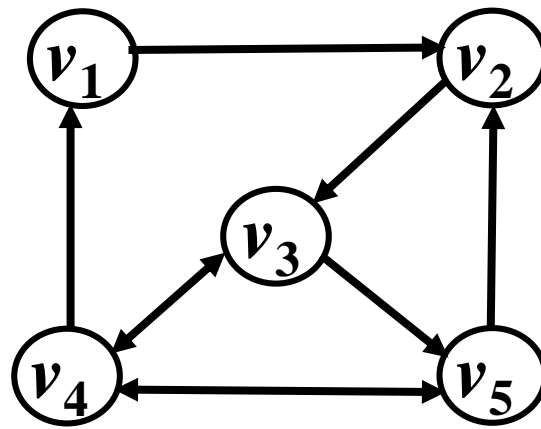
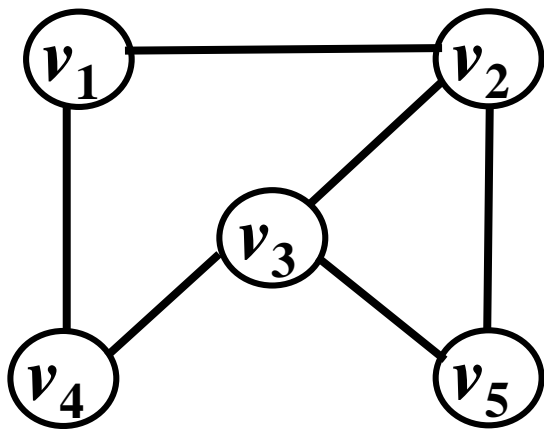
## 4.2 图的存储结构

### 是否可以采用顺序存储结构存储图(一维数组)？

- 图的特点：顶点之间的关系是 $m:n$ ，即任何两个顶点之间都可能存在关系（边），无法通过存储位置表示这种任意的逻辑关系，所以，图无法采用顺序存储结构。

### 如何存储图？

- 考虑图的定义，图是由顶点和边组成的；
- 如何存储**顶点**、如何存储**边**----顶点之间的关系。





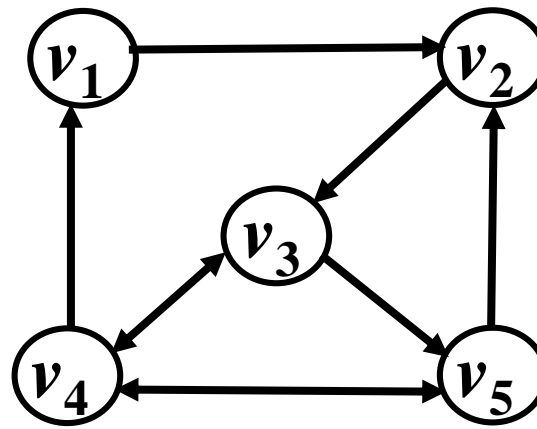
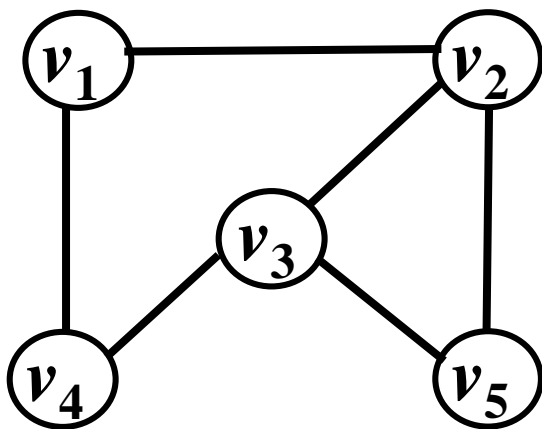
## 4.2 图的存储结构(cont.)

邻接矩阵 (*Adjacency Matrix*) 表示 (数组表示法)

➡ 基本思想:

- 用一个一维数组存储图中顶点的信息, 用一个二维数组 (称为邻接矩阵) 存储图中各顶点之间的邻接关系。
- 假设图  $G=(V, E)$  有  $n$  个顶点, 则邻接矩阵是一个  $n \times n$  的方阵, 定义为:

$$\text{edge}[i][j] = \begin{cases} 1 & \text{若 } (i, j) \in E \text{ 或 } \langle i, j \rangle \in E \\ 0 & \text{否则} \end{cases}$$

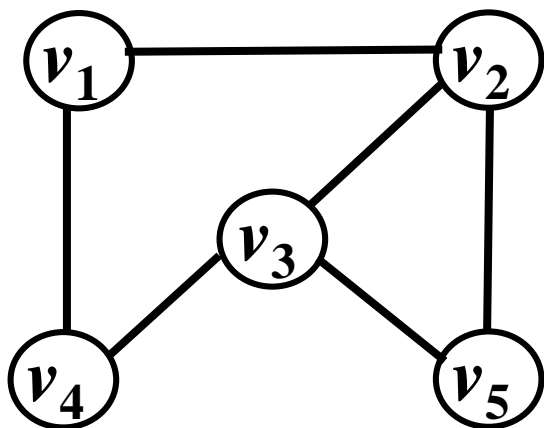




## 4.2 图的存储结构(cont.)

邻接矩阵 (*Adjacency Matrix*) 表示 (数组表示法)

➡ 无向图的邻接矩阵:



vertex =

	$v_1$	$v_2$	$v_3$	$v_4$	$v_5$
$v_1$	0	1	0	1	0
$v_2$	1	0	1	0	1
$v_3$	0	1	0	1	1
$v_4$	1	0	1	0	0
$v_5$	0	1	1	0	0

edge =

➡ 存储结构特点:

■ 主对角线为 0 且一定是对称矩阵;

问题: 1. 如何求顶点  $v_i$  的度?

2. 如何判断顶点  $v_i$  和  $v_j$  之间是否存在边?

3. 如何求顶点  $v_i$  的所有邻接点?

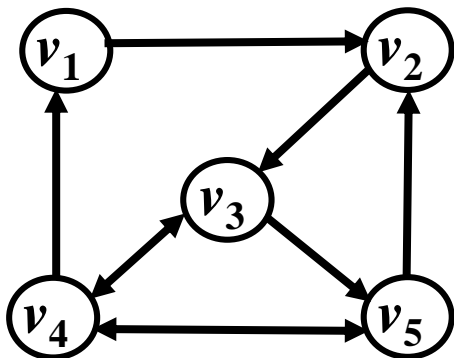




## 4.2 图的存储结构(cont.)

邻接矩阵 (Adjacency Matrix) 表示 (数组表示法)

➡ 有向图的邻接矩阵:



vertex =

	$v_1$	$v_2$	$v_3$	$v_4$	$v_5$
$v_1$	0	1	0	0	0
$v_2$	0	0	1	0	0
$v_3$	0	0	0	1	1
$v_4$	1	0	1	0	1
$v_5$	0	1	0	1	0

edge =

➡ 存储结构特点:

■ 有向图的邻接矩阵一定不对称吗?

问题: 1. 如何求顶点  $v_i$  的出度?

2. 如何判断顶点  $v_i$  和  $v_j$  之间是否存在有向边?

3. 如何求邻接于顶点  $v_i$  的所有顶点?

4. 如何求邻接到顶点  $v_i$  的所有顶点?





## 4.2 图的存储结构(cont.)

邻接矩阵 (*Adjacency Matrix*) 表示 (数组表示法)

➡ 存储结构定义:

假设图  $G$  有  $n$  个顶点  $e$  条边, 则该图的存储需求为

$O(n+n^2) = O(n^2)$ , 与边的条数  $e$  无关。

typedef struct {

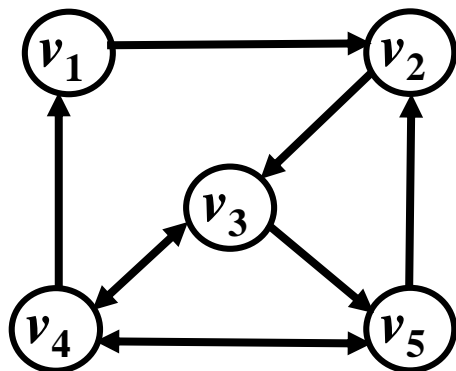
VertexData verlist [NumVertices]; // 顶点表

EdgeData edge[NumVertices][NumVertices];

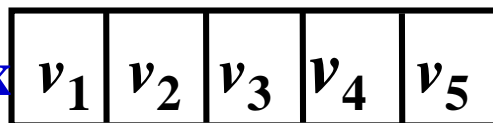
// 邻接矩阵—边表, 可视为边之间的关系

int n, e; // 图的顶点数与边数

} MTGraph;



vertex



edge=

$v_1$	$v_2$	$v_3$	$v_4$	$v_5$	
0	1	0	0	0	$v_1$
0	0	1	0	0	$v_2$
0	0	0	1	1	$v_3$
1	0	1	0	1	$v_4$
0	1	0	1	0	$v_5$

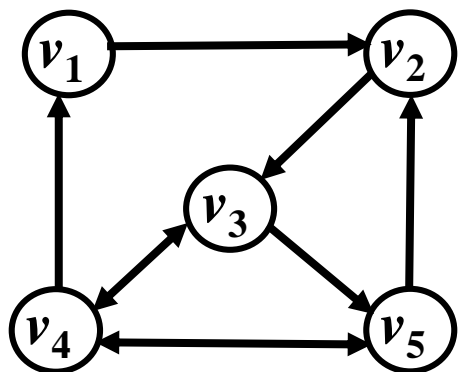




## 4.2 图的存储结构(cont.)

➡ 存储结构的建立----算法实现的步骤:

1. 确定图的顶点个数 $n$ 和边数 $e$ ;
2. 输入顶点信息存储在一维数组 $vertex$ 中;
3. 初始化邻接矩阵;
4. 依次输入每条边存储在邻接矩阵 $edge$ 中;
  - 4.1 输入边依附的两个顶点的序号 $i, j$ ;
  - 4.2 将邻接矩阵的第 $i$ 行第 $j$ 列的元素值置为1;
  - 4.3 将邻接矩阵的第 $j$ 行第 $i$ 列的元素值置为1。



$vertex$

$v_1$	$v_2$	$v_3$	$v_4$	$v_5$
$v_1$	$v_2$	$v_3$	$v_4$	$v_5$

$edge =$

$v_1$	$v_2$	$v_3$	$v_4$	$v_5$	
0	1	0	0	0	$v_1$
0	0	1	0	0	$v_2$
0	0	0	1	1	$v_3$
1	0	1	0	1	$v_4$
0	1	0	1	0	$v_5$





## 4.2 图的存储结构(cont.)

➤ 存储结构的建立算法的实现:

```
void CreateMGraph (MTGraph *G) //建立图的邻接矩阵
{
    int i, j, k, w;
    cin >> G->n >> G->e;           //1.输入顶点数和边数
    for (i=0; i<G->n; i++)           //2.读入顶点信息，建立顶点表
        G->verlist[i]=getchar();
    for (i=0; i<G->n; i++)
        for (j=0; j<G->n; j++)
            G->edge[i][j]=0;         //3.邻接矩阵初始化
    for (k=0; k<G->e; k++) {         //4.读入e条边建立邻接矩阵
        cin >> i >> j >> w;         // 输入边 (i,j) 上的权值w
        G->edge[i][j]=w; G->edge[j][i]=w;
    }
} //时间复杂度:  $T = O(n + n^2 + e)$ 。当  $e \ll n$ ,  $T = O(n^2)$  ?
```





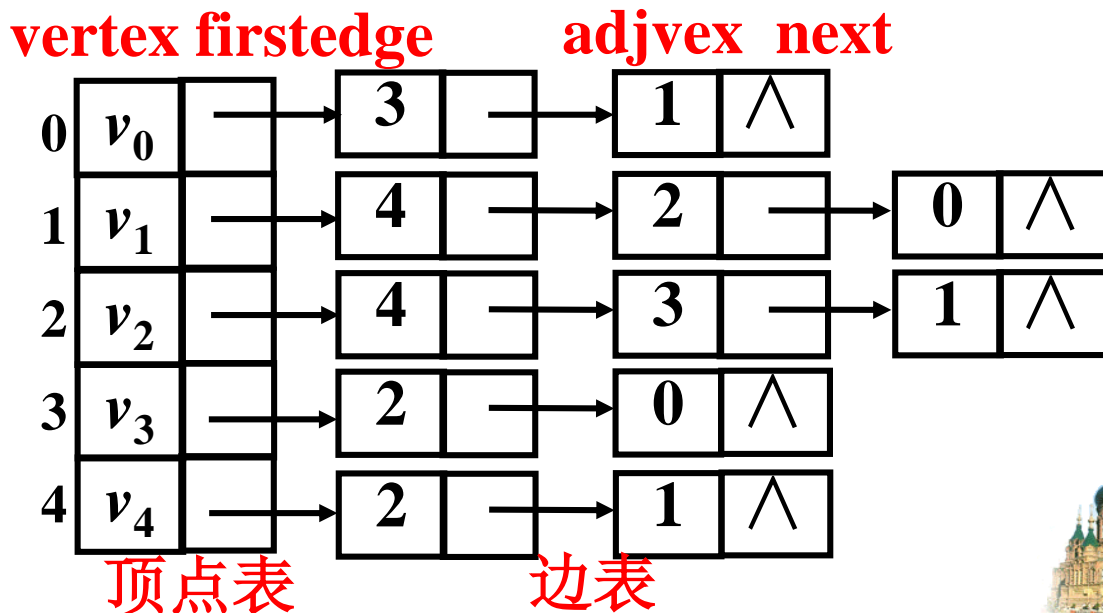
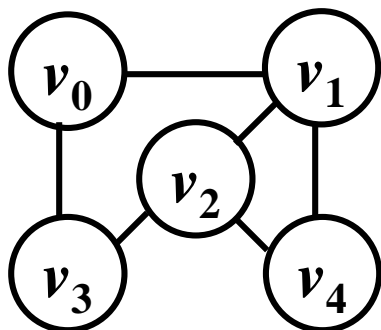


## 4.2 图的存储结构(cont.)

邻接表 (Adjacency List) 表示

➡ 无向图的邻接表:

- 对于无向图的每个顶点 $v_i$ ，将所有与 $v_i$ 相邻的顶点链成一个单链表，称为顶点 $v_i$ 的边表（顶点 $v_i$ 的邻接表）；
- 再把所有边表的指针和存储顶点信息的一维数组构成顶点表。

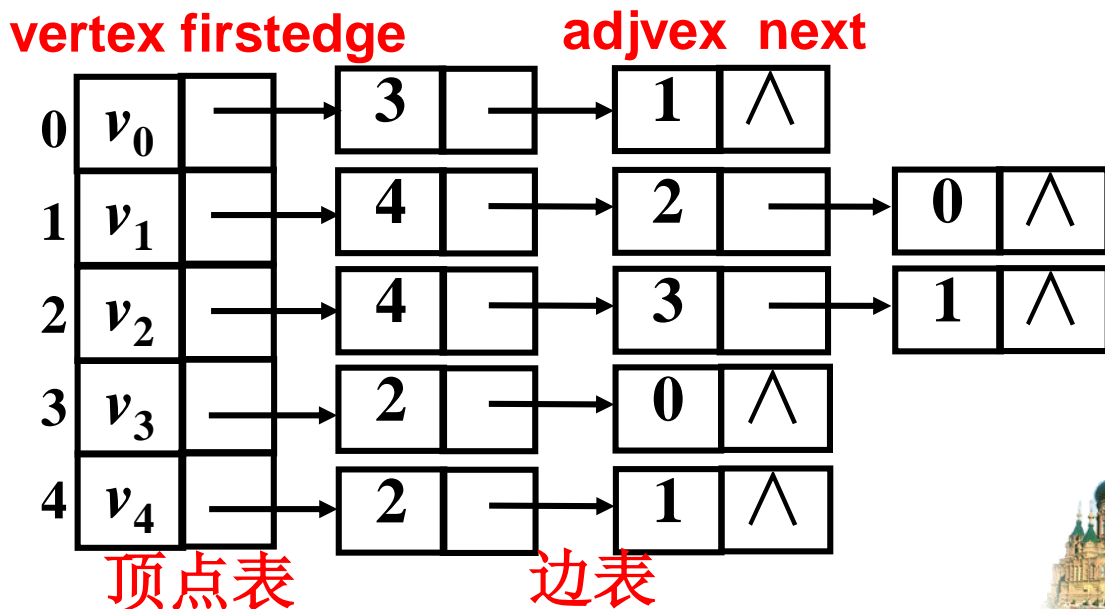
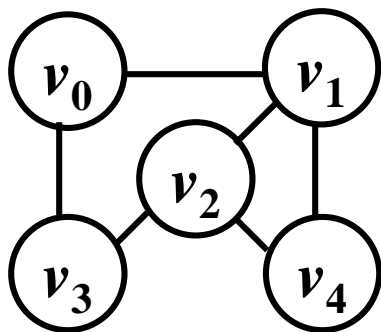




## 4.2 图的存储结构(cont.)

### ➤ 无向图的邻接表存储的特点:

- 边表中的结点表示什么?
- 如何求顶点  $v_i$  的度?
- 如何判断顶点  $v_i$  和顶点  $v_j$  之间是否存在边?
- 如何求顶点  $v_i$  的所有邻接点?
- 空间需求  $O(n+2e)$



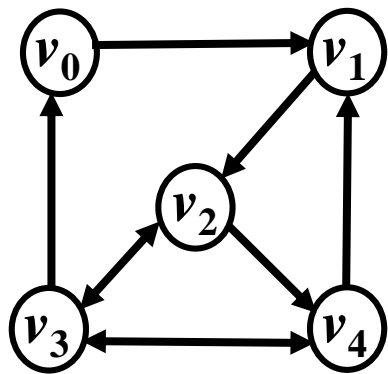


## 4.2 图的存储结构(cont.)

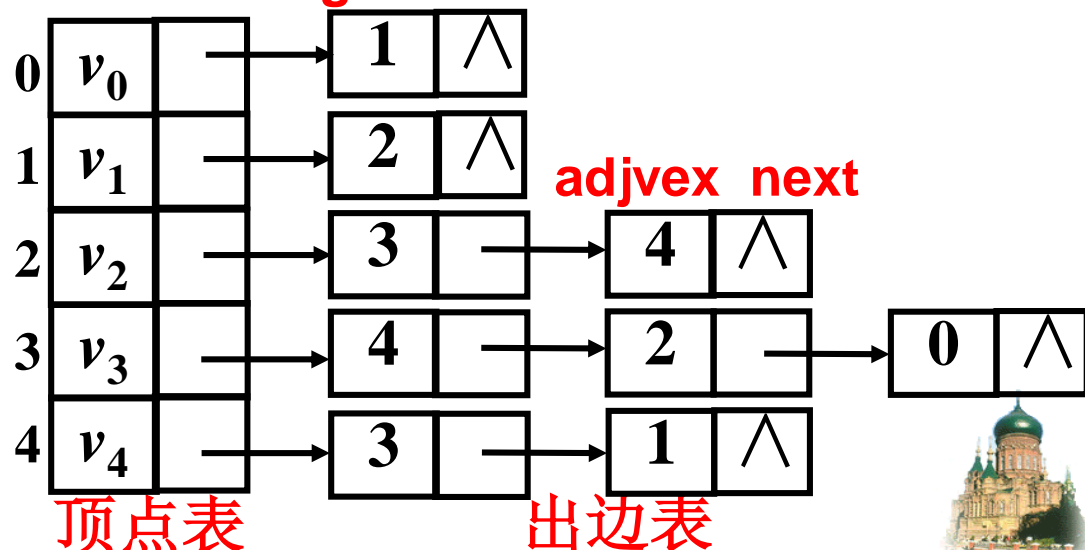
邻接表 (Adjacency List) 表示

➔ 有向图的邻接表——正邻接表

- 对于有向图的每个顶点 $v_i$ ，将邻接于 $v_i$ 的所有顶点链成一个单链表，称为顶点 $v_i$ 的出边表；
- 再把所有出边表的指针和存储顶点信息的一维数组构成顶点表。



vertex firstedge

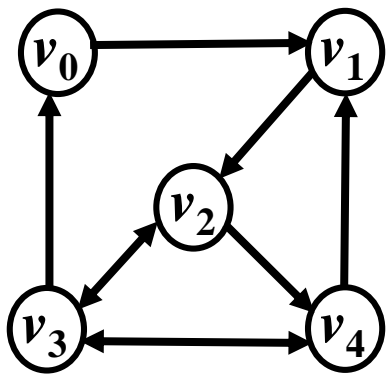




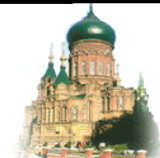
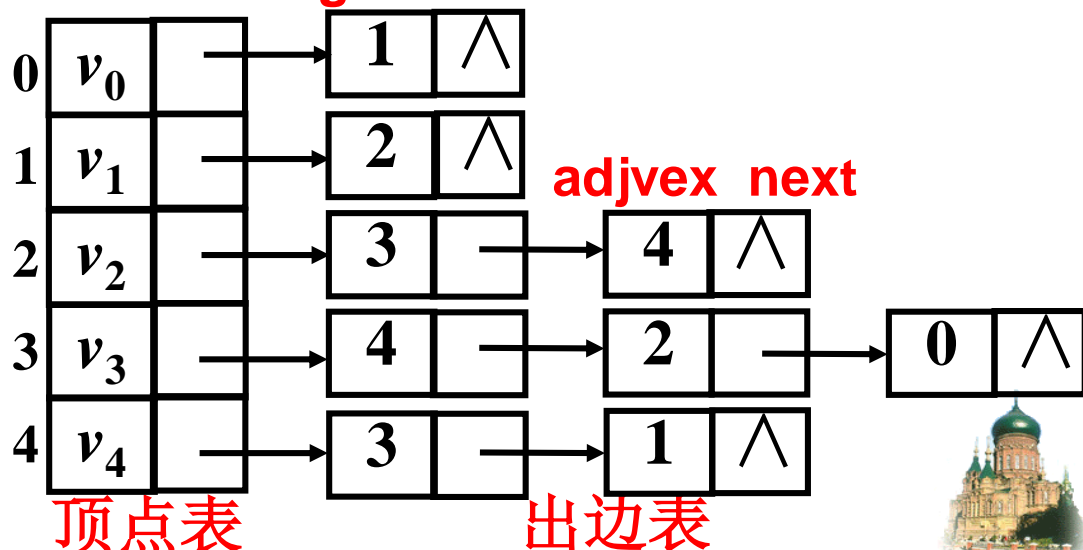
## 4.2 图的存储结构(cont.)

### 有向图的正邻接表的存储特点

- 出边表中的结点表示什么？
- 如何求顶点  $v_i$  的出度？如何求顶点  $v_i$  的入度？
- 如何判断顶点  $v_i$  和顶点  $v_j$  之间是否存在有向边？
- 如何求邻接于顶点  $v_i$  的所有顶点？
- 如何求邻接到顶点  $v_i$  的所有顶点？
- 空间需求:  $O(n+e)$



vertex firstedge



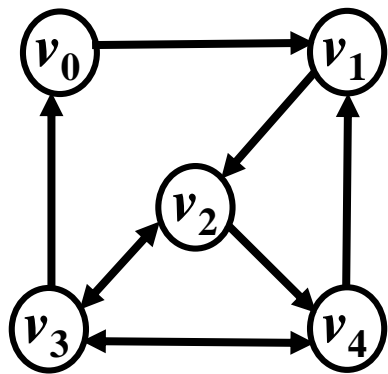


## 4.2 图的存储结构(cont.)

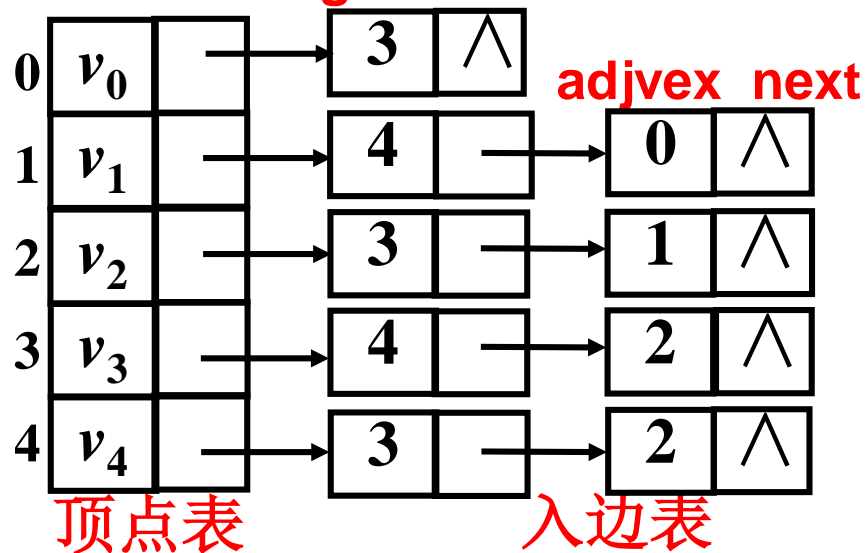
邻接表 (Adjacency List) 表示

➡ 有向图的邻接表——逆邻接表

- 对于有向图的每个顶点 $v_i$ ，将邻接到 $v_i$ 的所有顶点链成一个单链表，称为顶点 $v_i$ 的入边表；
- 再把所有入边表的指针和存储顶点信息的一维数组构成顶点表。



vertex firstedge

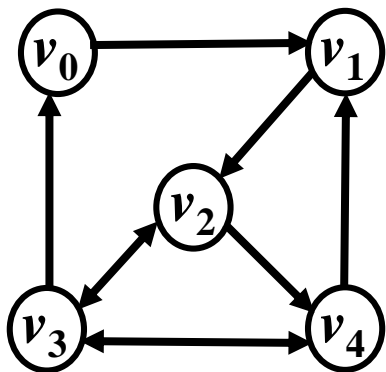




## 4.2 图的存储结构(cont.)

### 有向图的逆邻接表的存储特点

- 出边表中的结点表示什么？
- 如何求顶点  $v_i$  的入度？如何求顶点  $v_i$  的出度？
- 如何判断顶点  $v_i$  和顶点  $v_j$  之间是否存在有向边？
- 如何求邻接到顶点  $v_i$  的所有顶点？
- 如何求邻接于顶点  $v_i$  的所有顶点？
- 空间需求:  $O(n+e)$



vertex firstedge

0	$v_0$	—	3	^	
1	$v_1$	—	4	—	0 ^
2	$v_2$	—	3	—	1 ^
3	$v_3$	—	4	—	2 ^
4	$v_4$	—	3	—	2 ^

顶点表

入边表





## 4.2 图的存储结构(cont.)

### 邻接表存储结构的定义

```

typedef struct node { //边表结点
    int adjvex;        //邻接点域（下标）
    EdgeData cost;     //边上的权值
    struct node *next; //下一边链接指针
} EdgeNode;

typedef struct { //顶点表结点
    VertexData vertex; //顶点数据域
    EdgeNode * firstedge; //边链表头指针
} VertexNode;

typedef struct { //图的邻接表
    VertexNode vexlist [NumVertices];
    int n, e;        //顶点个数与边数
} AdjGraph;
  
```

边表结点

adjvex	cost	next
--------	------	------

顶点表结点

vertex	firstedge
--------	-----------

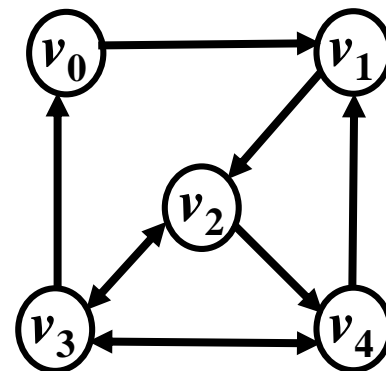




## 4.2 图的存储结构(cont.)

### ➤ 邻接表存储结构建立算法实现的步骤:

1. 确定图的顶点个数和边的个数;
2. 建立顶点表:
  - 2.1 输入顶点信息;
  - 2.2 初始化该顶点的边表;
3. 依次输入边的信息并存储在边表中;
  - 3.1 输入边所依附的两个顶点的序号tail和head和权值w;
  - 3.2 生成邻接点序号为head的边表结点p;
  - 3.3 设置边表结点p;
  - 3.4 将结点p插入到第tail个边表的头部;







## 4.2 图的存储结构(cont.)

➤ 邻接表存储结构建立算法的实现:

```
void CreateGraph (AdjGraph G)
```

```
{ cin >> G.n >> G.e;
```

```
  for ( int i = 0; i < G.n; i++) {
```

```
    cin >> G.vexlist[i].vertex;
```

```
    G.vexlist[i].firstedge = NULL; }
```

```
  for ( i = 0; i < G.e; i++) {
```

```
    cin >> tail >> head >> weight;
```

```
    EdgeNode * p = new EdgeNode;
```

```
    p->adjvex = head; p->cost = weight;
```

```
    p->next = G.vexlist[tail].firstedge;
```

```
    G.vexlist[tail].firstedge = p;
```

```
    p = new EdgeNode;
```

```
    p->adjvex = tail; p->cost = weight;
```

```
    p->next = G.vexlist[head].firstedge; //链入第 head 号链表的前端
```

```
    G.vexlist[head].firstedge = p; }
```

```
} //时间复杂度:  $O(2e+n)$ 
```

//1.输入顶点个数和边数

//2.建立顶点表

//2.1输入顶点信息

//2.2边表置为空表

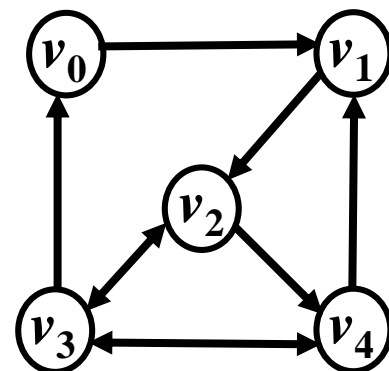
//3.逐条边输入,建立边表

//3.1输入

//3.2建立边结点

//3.3设置边结点

//3.4链入第 tail 号链表的前端





## 4.2 图的存储结构(cont.)

### 图的存储结构的比较——邻接矩阵和邻接表

	空间性能	时间性能	适用范围	唯一性
邻接矩阵	$O(n^2)$	$O(n^2)$	稠密图	唯一 ?
邻接表	$O(n+e)$	$O(n+e)$	稀疏图	不唯一 ?





## ◆ 十字链表(有向图)

- 十字链表是**有向图**的另一种**链式**存储结构。
- 将有向图的邻接表和逆邻接表结合起来的结构。
- 在十字链表中有两种结点：
  - ◆ **弧结点**：存储**每一条弧**的信息，用**链表**链接在一起。

弧结点结构：

<b>ivex</b>	<b>jvex</b>	<b>jlink</b>	<b>ilink</b>	<b>info</b>
-------------	-------------	--------------	--------------	-------------

- ◆ **jlink**:指向j的边； **ilink**:i发出的边

- ◆ **顶点结点**：存储**每一个顶点**的信息，使用**一维数组**来存储。

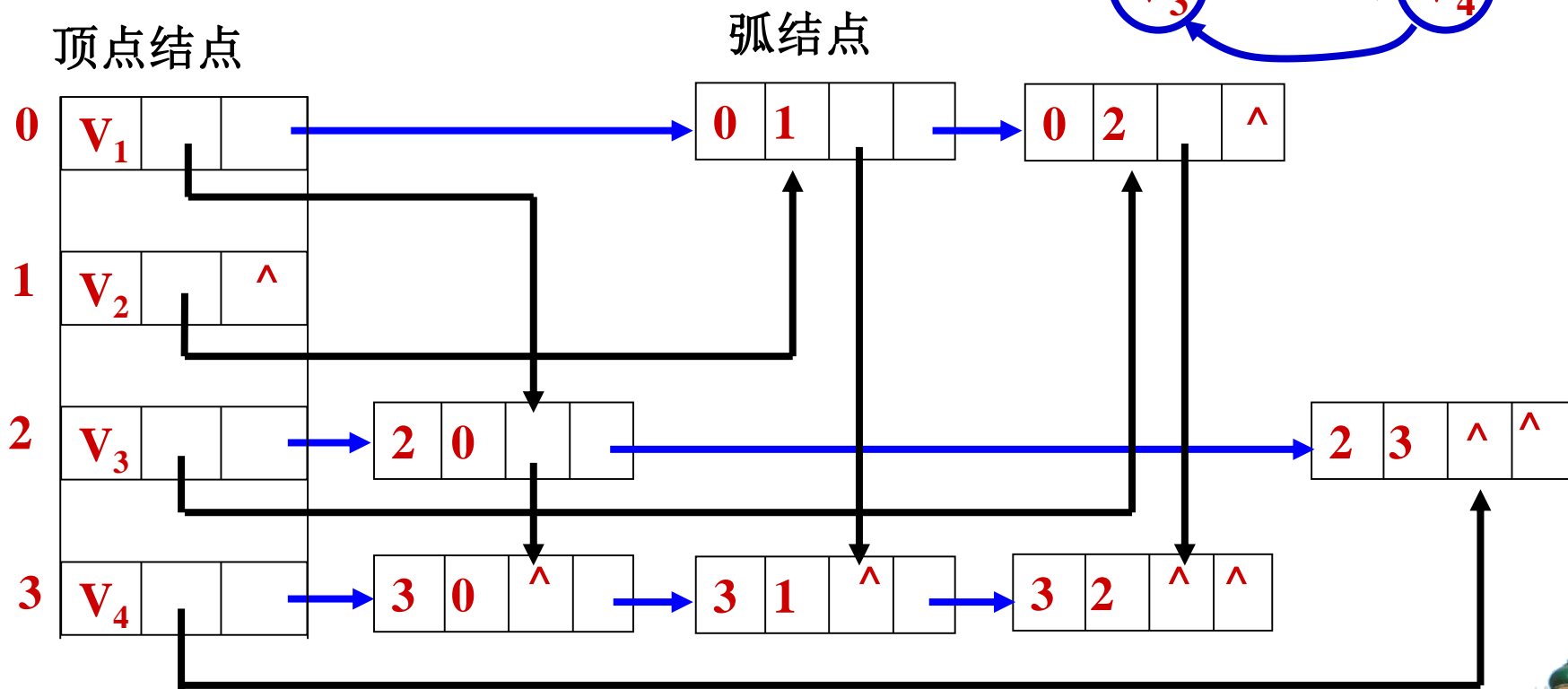
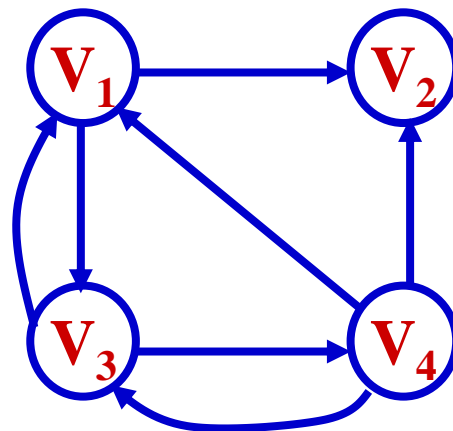
顶点结点结构：

<b>data</b>	<b>firstin</b>	<b>firstout</b>
-------------	----------------	-----------------



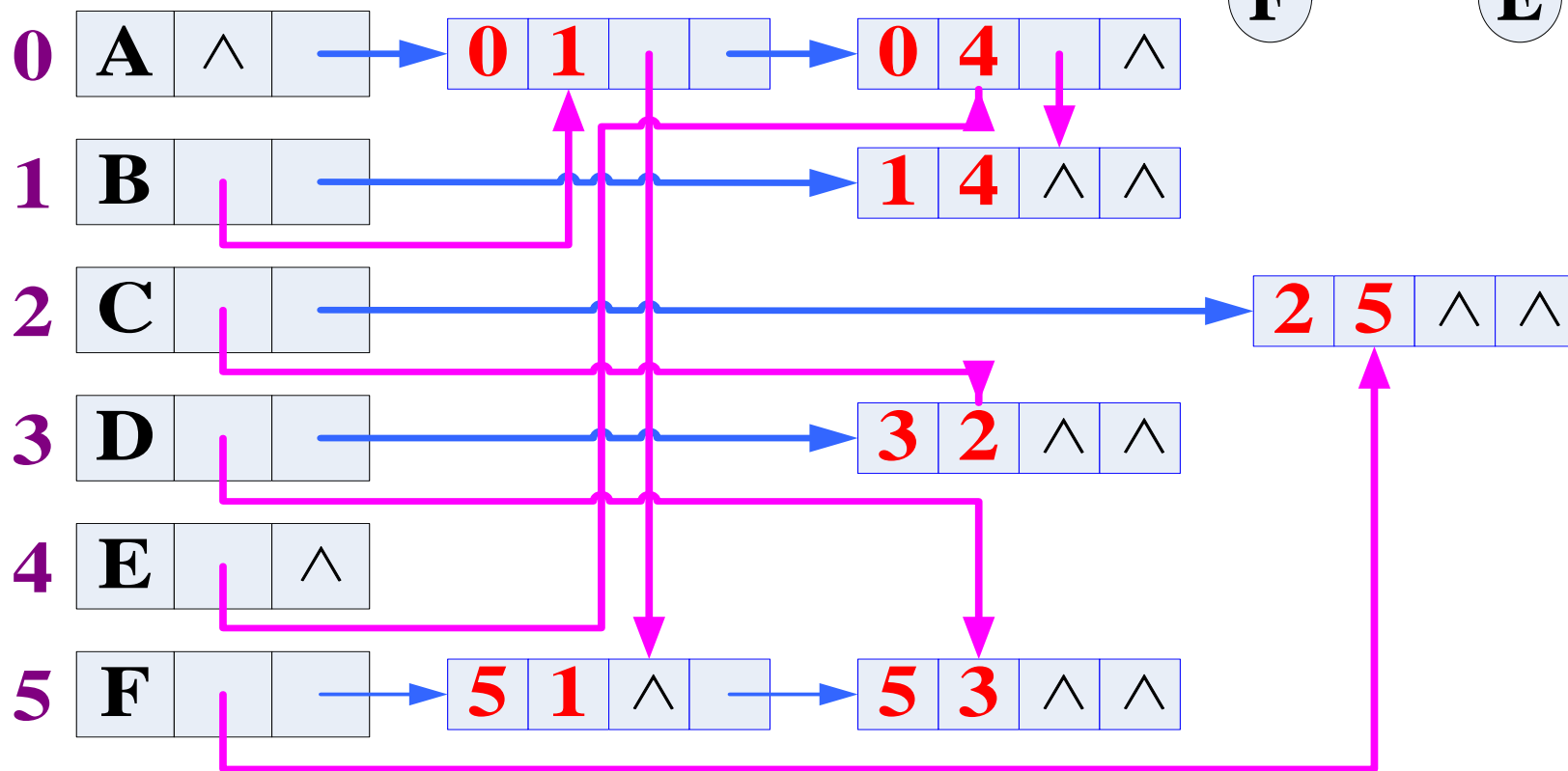
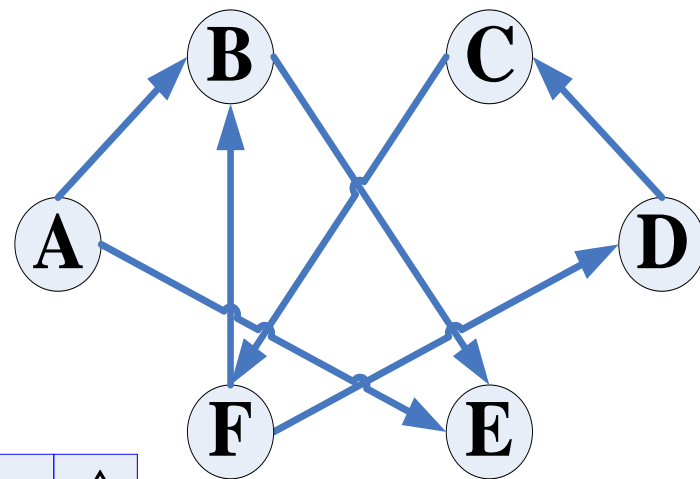


ivex	jvex	jlink	ilink	info
data	firstin	firstout		





➤ 十字链表中既容易找到以 $v_i$ 为尾的弧，也容易找到以 $v_i$ 为头的弧，因而容易求得顶点的出度和入度。





## ◆ 邻接多重表(无向图)

- 邻接多重表是无向图的另一种链式表示结构。
- 和十字链表类似。邻接多重表中，每一条边用一个结点表示。
- 在邻接多重表中有两种结点：
  - ◆ 边结点：存储每一条边的信息，用链表链接在一起。

边结点结构：

mark	ivex	ilink	jvex	jlink	info
------	------	-------	------	-------	------

- ◆ mark: 标识边是否被搜索过
- ◆ 顶点结点：存储每一个顶点的信息，使用一维数组来存储。

顶点结点结构：

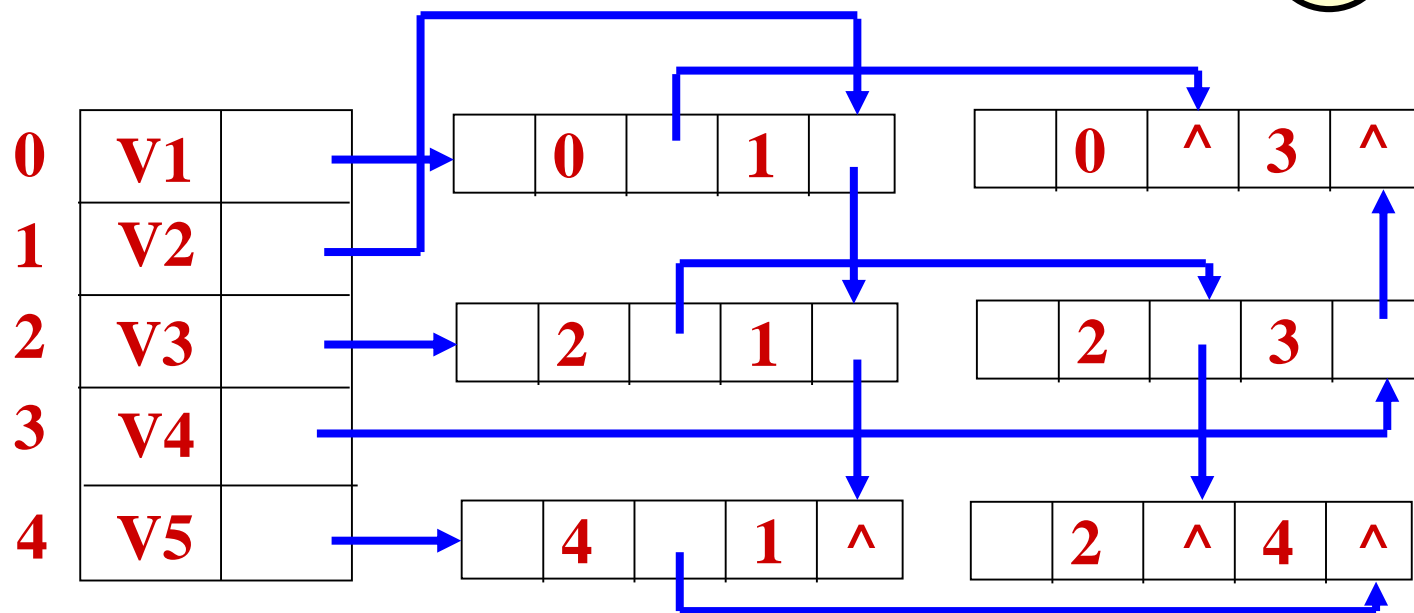
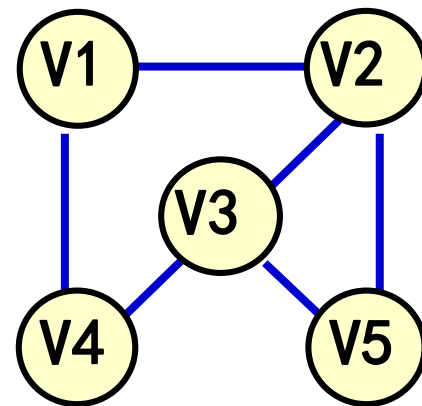
data	firstedge
------	-----------

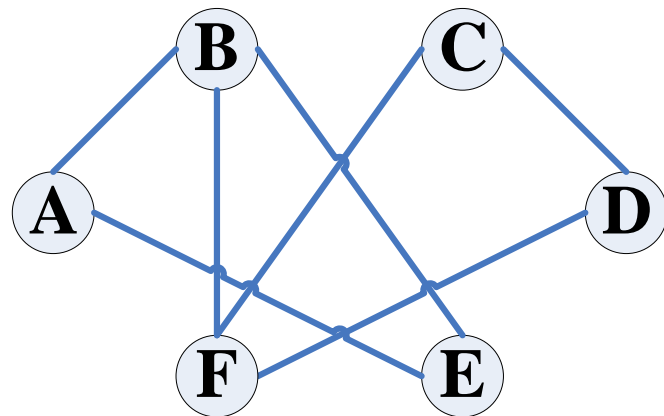
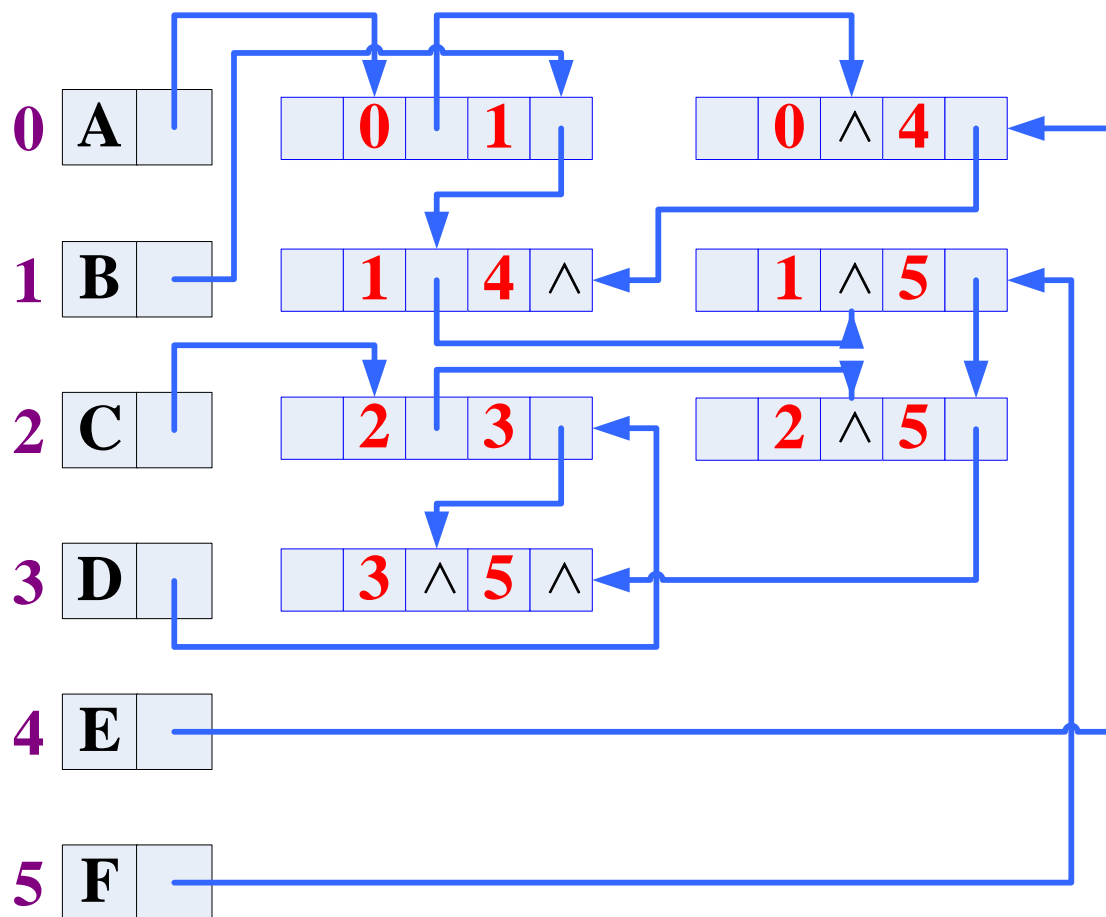




mark	ivex	ilink	jvex	jlink	info
------	------	-------	------	-------	------

data	firstedge
------	-----------









## 4.3 图的搜索（遍历）



**John Edward Hopcroft**

**Robert Endre Tarjan**



### 1986年图灵奖获得者

约翰·霍普克洛夫特1939年生于西雅图。美国国家科学院和工程院院士、康奈尔大学智能机器人实验室主任。1962和1964年获斯坦福大学硕士和博士学位。先后在普林斯顿大学、斯坦福大学等工作，也曾任职于一些科学研究机构如NSF和NRC。著作很多如《算法设计与分析基础》《数据结构与算法》《**自动机理论、语言和计算导论**》《形式语言及其与自动机的关系》

——  
 罗伯特·塔扬普林斯顿大学计算机科学系教授，1948年4月30日生于加利福尼亚州。1969年本科毕业，进入斯坦福大学研究生院，1972年获得博士学位。**平面图测试的高效算法**；合并-搜索问题；“分摊”算法的概念；八字形树；持久性数据结构





## 4.3 图的搜索（遍历） (cont.)

### ➤ 图的遍历（图的搜索）

- 从图中**某**一顶点出发，对图中所有顶点访问一次且仅**访问**一次。
- **访问**：抽象操作，可以是对结点进行的各种处理

### ➤ 图结构的复杂性

- 在**线性表**中，数据元素在表中的编号就是元素在序列中的位置，因而其**编号是唯一的**；
- 在**树结构**中，将结点按层序编号，由于树具有层次性，因而其**层序编号也是唯一的**；
- 在**图结构**中，任何两个顶点之间都可能存在边，顶点是没有确定的先后次序的，所以，**顶点的编号不唯一**。





## 4.3 图的搜索（遍历）(cont.)

### ➡ 图的遍历要解决的关键问题

- 在图中，如何选取遍历的起始顶点？
  - 解决办法：从编号小的顶点(任取一顶点，适合编程)开始。
- 从某个起点始可能到达不了所有其它顶点，怎么办？
  - 解决办法：多次调用遍历图（起点选没有用过的）的算法。
- 图中可能存在回路，且图的任一顶点都可能与其它顶点“相通”，在访问完某个顶点之后可能会沿着某些边又回到了曾经访问过的顶点。如何避免某些顶点可能会被重复访问？
  - 解决办法：设访问标志数组visited[n]。
- 在图中，一个顶点可以和其它多个顶点相连，当这样的顶点访问过后，如何选取下一个要访问的顶点？
  - 解决办法：**深度优先搜索**（Depth First Search）和**广度优先搜索**（Breadth First Search）。





## 4.3 图的搜索（遍历） (cont.)

➤ **深度优先遍历**----类似于树结构的先序遍历

设**图G**的**初态**是所有顶点都“未访问过（**False**）”，在**G**中任选一个顶点 **v** 为初始出发点(**源点**)，则**深度优先搜索**可**定义**为：

- ①首先访问出发点 **v**，并将其标记为“访问过（**True**）”；
- ②然后，从 **v** 出发，依次考察与 **v** 相邻的顶点 **w**；若 **w** “未访问过（**False**）”，则以 **w** 为新的出发点**递归地**进行**深度优先**搜索，直到图中所有与源点 **v** 有路径相通的顶点（亦称从源点可到达的顶点）均被访问为止；
- ③若此时图中仍有未被访问过的顶点，则另选一个“未访问过”的顶点作为新的搜索起点，重复上述过程，直到图中所有顶点都被访问过为止。

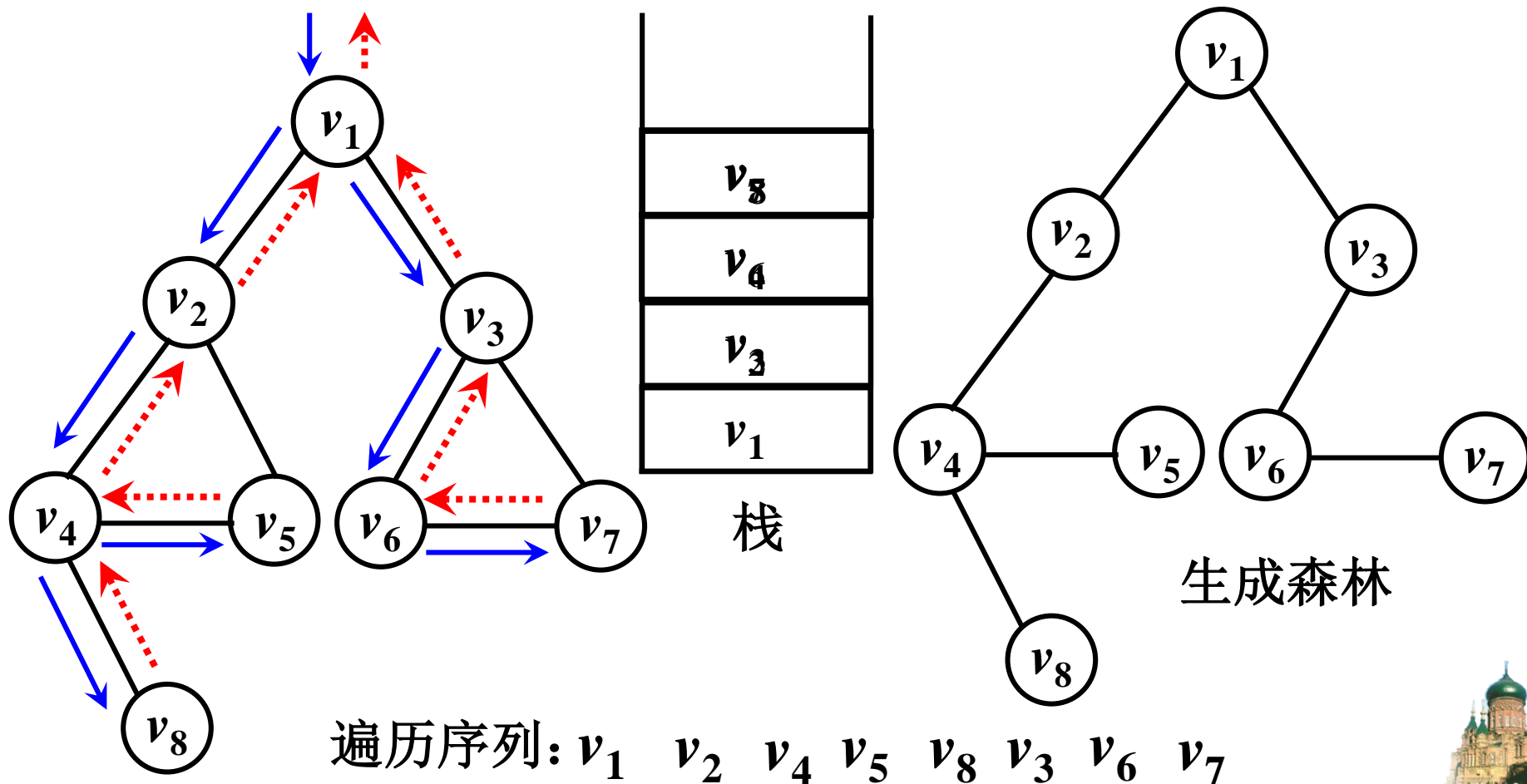




## 4.3 图的搜索（遍历）(cont.)

### 深度优先遍历示例

深度优先遍历序列? 入栈序列? 出栈序列?





## 4.3 图的搜索（遍历） (cont.)

### ➤ 深度优先遍历特点：

- 是递归的定义，是尽可能对纵深方向上进行搜索，故称**先深或深度优先搜索**。

### ➤ 先深或深度优先编号。

- 搜索过程中，根据访问顺序给顶点进行的编号，称为**先深或深度优先编号**。

### ➤ 先深序列或DFS序列：

- 先深搜索过程中，根据访问顺序得到的顶点序列，称为**先深序列或DFS序列**。

### ➤ 生成森林（树）：

- 有原图的**所有顶点**和搜索过程中所**经过的边**构成的子图。

### ➤ 先深搜索结果不唯一

- 即图的**DFS序列**、**先深编号**和**生成森林**不唯一。





## 4.3 图的搜索（遍历） (cont.)

➡ 深度优先遍历主算法:

`bool visited[NumVertices];` //访问标记数组是全局变量

`int dfn[NumVertices];` //顶点的先深编号

`void DFSTraverse (AdjGraph G)` //主算法

// 先深搜索----邻接表表示的图G; 而以邻接矩阵表示G时, 算法完全相同

```
{ int i, count = 1;
  for ( int i = 0; i < G.n; i++ )
    visited [i] =False; //标志数组初始化
  for ( int i = 0; i < G.n; i++ )
    if ( ! visited[i] )
      DFSX ( G, i ); //从顶点 i 出发的一次搜索, BFSX(G, i)
}
```







## 4.3 图的搜索（遍历）(cont.)

➡ 从一个顶点出发的一次深度优先遍历算法：

■ 实现步骤：

1. 访问顶点 $v$ ;  $visited[v]=1$ ;
2.  $w$ =顶点 $v$ 的第一个邻接点;
3. while ( $w$ 存在)
  - 3.1 if ( $w$ 未被访问) 从顶点 $w$ 出发递归执行该算法;
  - 3.2  $w$ =顶点 $v$ 的下一个邻接点;







## 4.3 图的搜索（遍历） (cont.)

➤ 从一个顶点出发的一次深度优先遍历算法:

```
void DFS1 (AdjGraph* G, int i)
```

//以 $v_i$ 为出发点时对邻接表表示的图G进行先深搜索

```
{   EdgeNode *p;
    cout<<G→vexlist[i].vertex;    //访问顶点 $v_i$ ;
    visited[i]=True;                //标记 $v_i$ 已访问
    dfn[i]=count++;                 //对 $v_i$ 进行编号
    p=G→vexlist[i].firstedge;      //取 $v_i$ 边表的头指针
    while( p ) { //依次搜索 $v_i$ 的邻接点 $v_j$ , 这里 $j=p→adjvex$ 
        if( !visited[ p→adjvex ] ) //若 $v_j$ 尚未访问
            DFS1(G, p→adjvex); //则以 $v_j$ 为出发点先深搜索
        p=p→next;
    }
} //DFS1
```





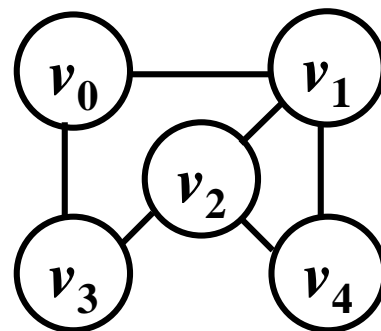
## 4.3 图的搜索（遍历） (cont.)

➡ 从一个顶点出发的一次深度优先遍历算法:

```
void DFS1 (AdjGraph* G, int i)
```

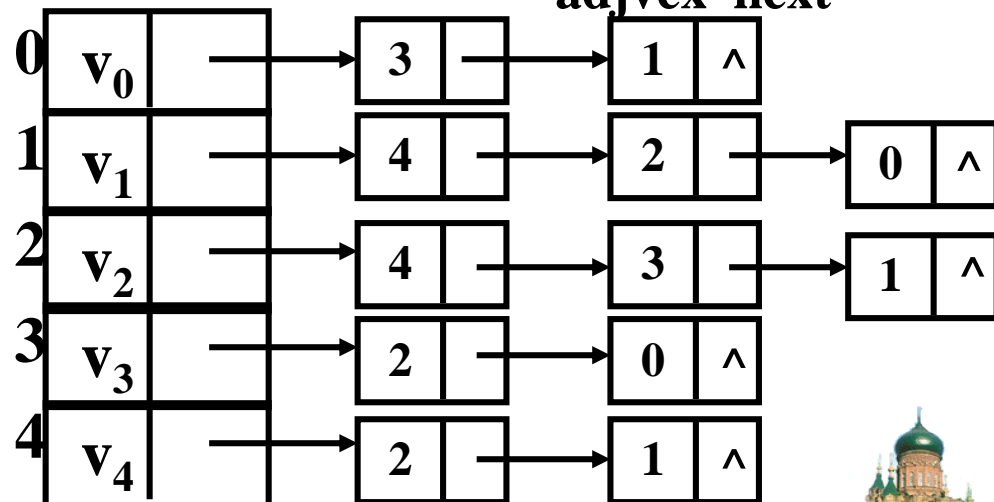
//以 $v_i$ 为出发点时对邻接表表示的图G进行先深搜索

```
{   EdgeNode *p;
    cout<<G→vexlist[i].vertex;
    visited[i]=True;
    dfn[i]=count++;
    p=G→vexlist[i].firstedge;
    while( p ) {
        if( !visited[ p→adjvex ] )
            DFS1(G, p→adjvex);
        p=p→next;
    } //DFS1
```



vertex firstedge

adjvex next



顶点表

边表





## 4.3 图的搜索（遍历） (cont.)

➔ 从一个顶点出发的一次深度优先遍历算法:

```
void DFS2(MTGraph *G, int i)
```

```
//以 $v_i$ 为出发点对邻接矩阵表示的图G进行深度优先搜索
```

```
{ int j;
```

```
    cout<<G→vexlist[i];    //访问定点 $v_i$ 
```

```
    visit[i]=True;          //标记 $v_i$ 已访问
```

```
    dfn[i]=count;           //对 $v_i$ 进行编号
```

```
    count ++;               //下一个顶点的编号
```

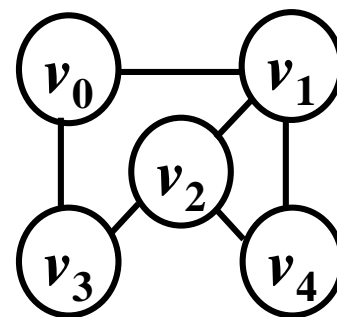
```
    for( j=0; j<G→n; j++ ) //依次搜索 $v_i$ 的邻接点
```

```
        if ( (G→edge[i][j] == 1)&& ! visited[j] ) //若 $v_j$ 尚未访问
```

```
            DFS2( G, j );
```

```
}//DFS2
```

	$v_0$	$v_1$	$v_2$	$v_3$	$v_4$
$v_0$	0	1	0	1	0
$v_1$	1	0	1	0	1
$v_2$	0	1	0	1	1
$v_3$	1	0	1	0	0
$v_4$	0	1	1	0	0





## 4.3 图的搜索（遍历） (cont.)

➔ **广度优先遍历**----类似于树结构的层序遍历

设**图G**的**初态**是所有顶点都“未访问过（False）”，在G中任选一个顶点  $v$  为**源点**，则**广度优先搜索**可**定义**为：

- ①首先访问出发点  $v$ ，并将其标记为“访问过（True）”；
- ②接着依次访问所有**与  $v$  相邻**的顶点  $w_1, w_2 \dots w_t$ ；
- ③然后依次访问**与  $w_1, w_2 \dots w_t$  相邻**的所有未访问的顶点；
- ④依次类推，直至图中所有与源点  $v$  有路相通的顶点都已访问过为止；
- ⑤此时，从  $v$  开始的搜索结束，若G是连通的，则遍历完成；否则在G中另选一个尚未访问的顶点作为新源点，继续上述搜索过程，直到G中的所有顶点均已访问为止。

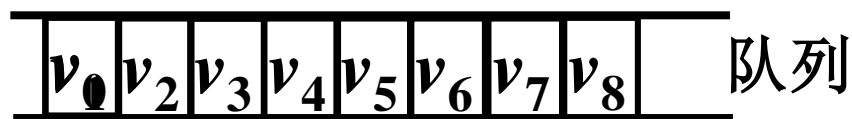
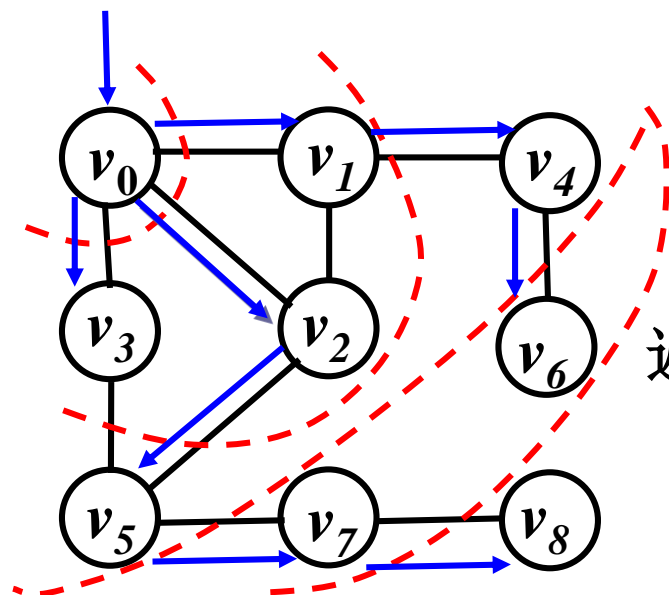




## 4.3 图的搜索（遍历） (cont.)

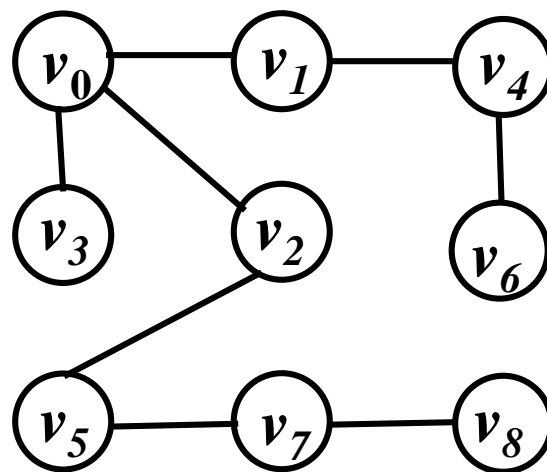
➤ 广度优先遍历示例

➤ 广度优先遍历序列? 入队序列? 出队序列?



遍历序列:  $v_0 \ v_1 \ v_2 \ v_3 \ v_4 \ v_5 \ v_6 \ v_7 \ v_8$

生成森林





## 4.3 图的搜索（遍历） (cont.)

### ➤ 广度优先遍历特点：

- 尽可能横向上进行搜索，并使“先被访问的顶点的邻接点”先于“后被访问的顶点的邻接点”被访问，故称**先广搜索**或**广度优先搜索**。

### ➤ 先广或广度优先编号：

- 搜索过程中，根据访问顺序给顶点进行的编号，称为**先广**或**广度优先编号**

### ➤ 先广序列或BFS序列：

- 先广搜索过程中，根据访问顺序得到的顶点序列，称为**先广序列**或**BFS序列**。

### ➤ 生成森林（树）：

- 有原图的**所有顶点**和搜索过程中所**经过的边**构成的子图。

### ➤ 先广搜索结果不唯一：

- 即图的**BFS序列**、**先广编号**和**生成森林**不唯一。





## 4.3 图的搜索（遍历） (cont.)

➡ 广度优先遍历主算法:

`bool visited[NumVertices];` //访问标记数组是全局变量

`int dfn[NumVertices];` //顶点的先深编号

`void BFSTraverse (AdjGraph G)` //主算法

*/\* 先广搜索一邻接表表示的图G；而以邻接矩阵表示G时，算法完全相同*

```
{ int i, count = 1;
```

```
  for ( int i = 0; i < G.n; i++ )
```

```
    visited [i] =False; //标志数组初始化
```

```
  for ( int i = 0; i < G.n; i++ )
```

```
    if ( ! visited[i] )
```

```
      BFSX ( G,i ); //从顶点 i 出发的一次搜索, DFSX (G,i)
```

```
}
```





## 4.3 图的搜索（遍历）(cont.)

➔ 从一个顶点出发的一次广度优先遍历算法：

■ 实现步骤：

1. 初始化队列Q;
2. 访问顶点v;  $\text{visited}[v]=1$ ; 顶点v入队Q;
3. while (队列Q非空)
  - 3.1 v=队列Q的队头元素出队;
  - 3.2 w=顶点v的第一个邻接点;
  - 3.3 while (w存在)
    - 3.3.1 如果w 未被访问，则  
访问顶点w;  $\text{visited}[w]=1$ ; 顶点w入队列Q;
    - 3.3.2 w=顶点v的下一个邻接点;







## 4.3 图的搜索（遍历） (cont.)

```

void BFS1 (AdjGraph *G, int k)//这里没有进行先广编号
{   int i; EdgeNode *p; Queue Q;  MakeNull(Q);
    cout << G→vexlist[ k ].vertex;  visited[ k ] = True;
    EnQueue (k, Q);                  //进队列
    while ( ! Empty (Q) ) {          //队空搜索结束
        i=DeQueue(Q);                //vi出队
        p =G→vexlist[ i ].firstedge; //取vi的边表头指针
        while ( p ) {                //若vi的邻接点 vj (j= p→adjvex)存在,依次搜索
            if ( !visited[ p→adjvex ] ) { //若vj未访问过
                cout << G→vexlist[ p→adjvex ].vertex; //访问vj
                visited[ p→adjvex ]=True;             //给vj作访问过标记
                EnQueue ( p→adjvex , Q );             //访问过的vj入队
            }
            p = p→next;                //找vi的下一个邻接点
        } // 重复检测 vi的所有邻接顶点
    } //外层循环，判队列空否
} //以vk为出发点时对用邻接表表示的图G进行先广搜索
  
```





## 4.3 图的搜索（遍历） (cont.)

```

void BFS2 (MTGraph *G, int k) //这里没有进行先广编号
{
    int i, j; Queue Q; MakeNull(Q);
    cout << G→vexlist[ k ]; //访问  $v_k$ 
    visited[ k ] = True; //给  $v_k$  作访问过标记
    EnQueue (k, Q); //  $v_k$  进队列
    while ( ! Empty (Q) ) { //队空时搜索结束
        i=DeQueue(Q); //  $v_i$  出队
        for(j=0; j<G→n; j++) { //依次搜索  $v_i$  的邻接点  $v_j$ 
            if ( G→edge[ i ][ j ] ==1 && !visited[ j ] ) { //若  $v_j$  未访问过
                cout << G→vexlist[ j ]; //访问  $v_j$ 
                visited[ j ]=True; //给  $v_j$  作访问过标记
                EnQueue ( j , Q ); //访问过的  $v_j$  入队
            }
        } //重复检测  $v_i$  的所有邻接顶点
    } //外层循环，判队列空否
} //以  $v_k$  为出发点时对用邻接矩阵表示的图G进行先广搜索
  
```





# 无向图（的搜索）及其应用

## ■ 无向图连通性判定

### ■ 不连通：若干个生成树

求连通分量个数；

求出每个连通分量；

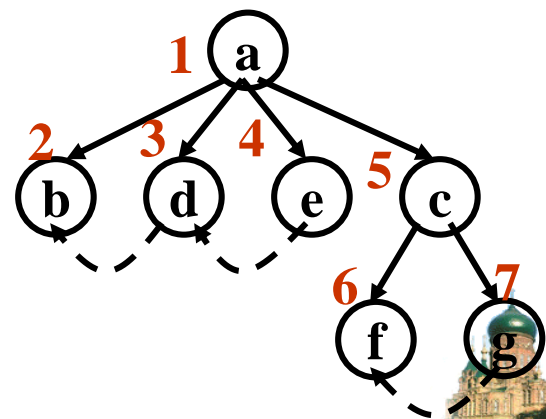
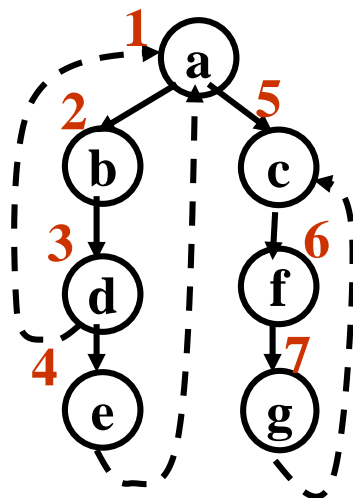
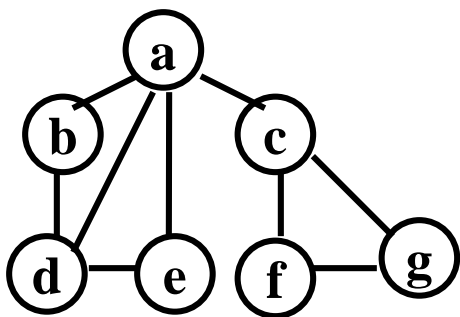
### ■ 连通：一棵生成树

判断是否有环路；

求带权连通图的最小生成树；4.3和4.4

判断是否是双连通的4.5

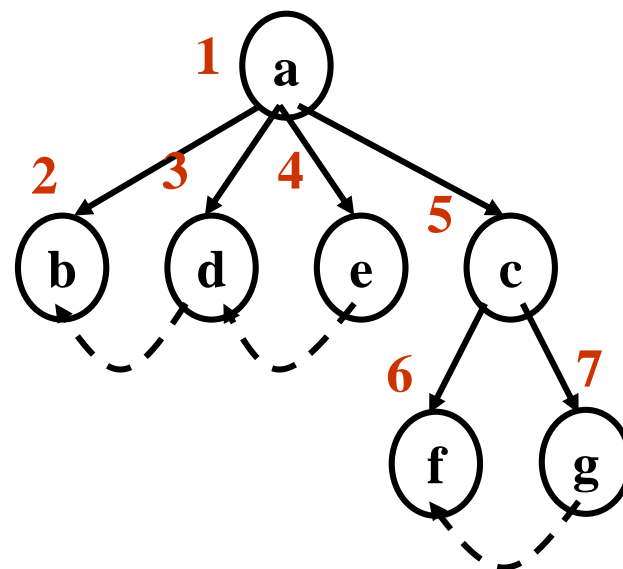
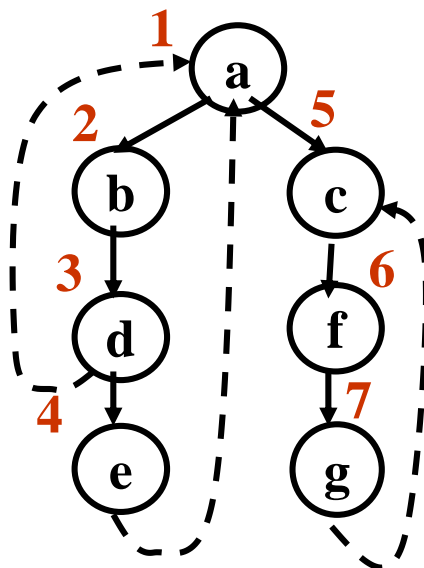
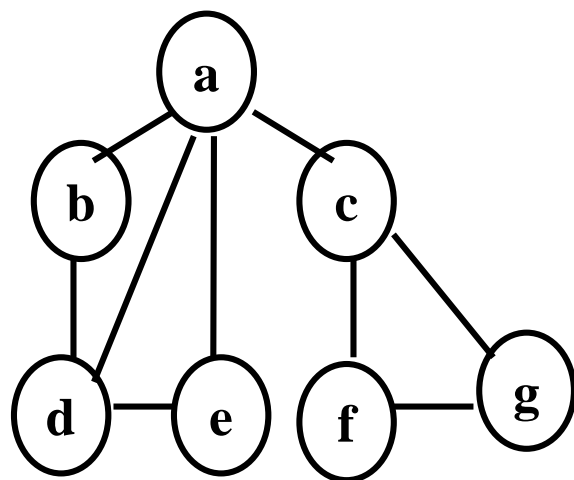
求关节点和双连通分量。





## 4.4 图与树的联系

### 4.4.1 先深生成森林和先广生成森林





## 一、搜索的结果

生成森林和（先深或先广）编号—顶点的线性序列。

树边	与	非树边
连通图		一个生成树
非连通图		生成森林
		连通子图（连通分量）

## 二、搜索过程中对边的分类

### 1. 先深搜索对边的分类

- 两类： 树边—在搜索过程中所经过的边；回退边—图中的其它边.
- 特点： 树边是从先深编号较小的指向较大的顶点；回退边相反；

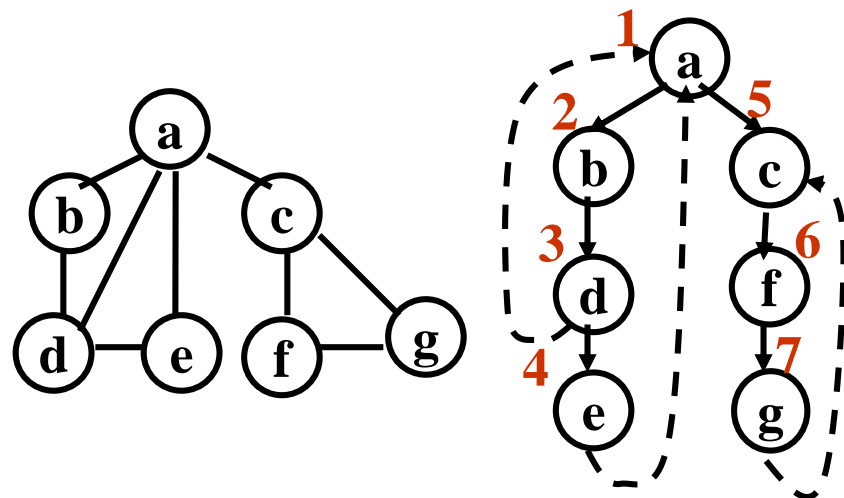




## 如何在搜索过程中区分树边和回退边？

设 $v$ 是当前访问过的顶点old，而下面搜索到 $w$ ，则 $w$ 有三种情况：

1.  $w$ 是new，则 $(v, w)$ 是树边，将其加入 $T$ ；
2.  $w$ 是old，且 $w$ 是 $v$ 的父亲，则 $(w, v)$ 是树边，但是第二次遇到，不再加入 $T$ ；例 $v=e, w=d$
3.  $w$ 是old且 $w$ 不是 $v$ 的父亲，则 $(v, w)$ 是回退边；例 $v=e, w=a$



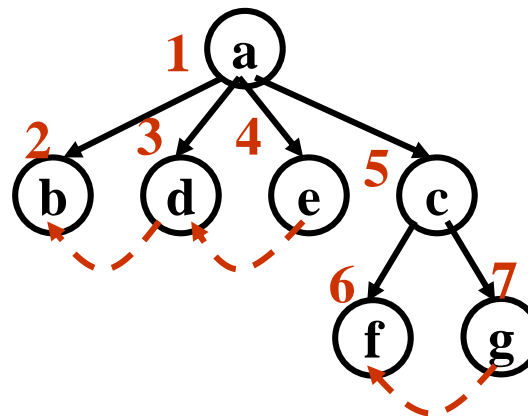
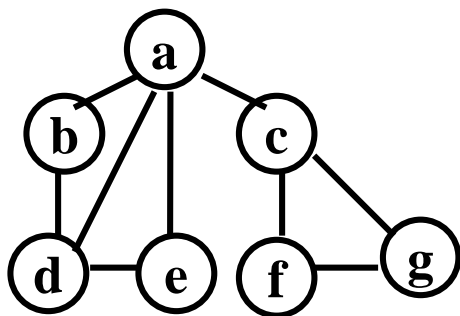
■ 结论：若 $G$ 中存在环路，则在先深搜索过程中必遇到回退边；反之亦然





## 2. 先广搜索对边的分类

- 两类：树边—在搜索过程中所经过的边；横边—图中的其它边。
- 特点：树边是从先深编号较小的指向较大的顶点；  
而横边不一定与之相反，但可规定：大→小。



- 结论：若G中存在环路，则在先广搜索过程中必遇到横边；反之亦然





## 4.4.2 无向图与开放树的联系

**定义：** 连通而无环路的无向图称作开放树 (Free Tree)

**开放树的性质：** (证明见教材P<sub>154~155</sub>)

- (1) 具有 $n \geq 1$ 个顶点的开放树包含 $n-1$ 条边;
- (2) 如果在开放树中任意加上一条边, 便得到一条回路。

**联系：** 如果指定开放树中的一个顶点为根(如搜索起点), 并且把每条边看成是背离根的, 则一株开放树变成一株树。







## 4.4.3 最小生成树算法

### 生成树的代价

- 设 $G=(V, E)$ 是一个无向连通网， $E$ 中每一条边 $(u, v)$ 上的权值 $c(u, v)$ ，称为 $(u, v)$ 的边长。
- 图 $G$ 的生成树上各边的权值（边长）之和称为该生成树的代价。

### 最小生成树(Minimum-Cost Spanning Tree, MST)

- 在图 $G$ 所有生成树中，代价最小的生成树称为最小生成树

### 最小生成树的概念可以应用到许多实际问题中。

- 例如，在 $n$ 个教室之间建造局域网络，至少要架设 $n-1$ 条通信线路，而每两个教室之间的距离可能不同，从而架设通信线路的造价就是不一样的，那么如何设计才能使得总造价最小？





### ◆ 构造最小生成树的准则

- 必须使用且仅使用该连通图中的 $n-1$ 条边连接结图中的 $n$ 个顶点;
- 不能使用产生回路的边;
- 各边上的权值的总和达到最小。

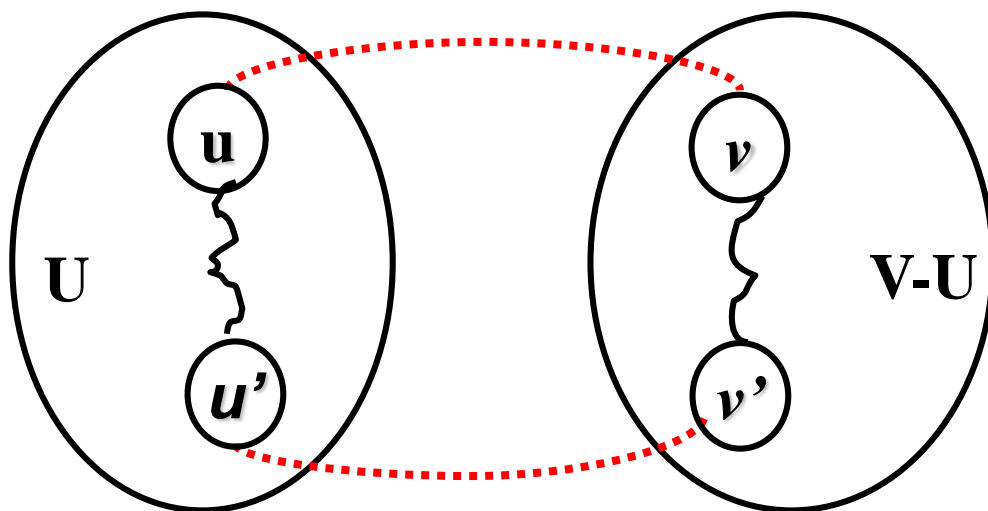




## 4.4 最小生成树算法(cont.)

### 最小生成树的性质

- 假设  $G = (V, E)$  是一个连通网， $U$  是顶点  $V$  的一个非空子集。若  $(u, v)$  是一条具有最小权值（代价）的边，其中  $u \in U$ ， $v \in V-U$ ，则必存在一棵包含边  $(u, v)$  的最小生成树。
- 此性质保证了Prim和Kruskal贪心算法的正确性

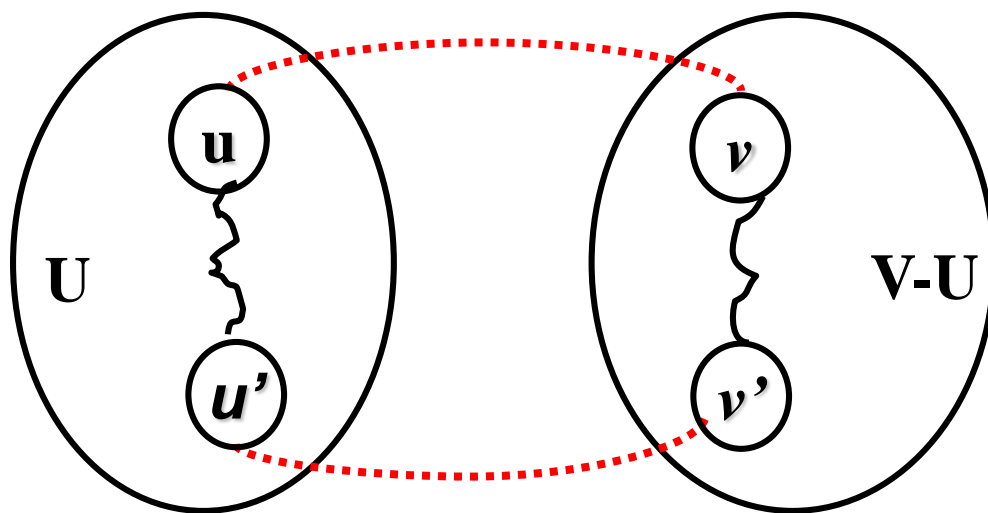




## 4.4 最小生成树算法(cont.)

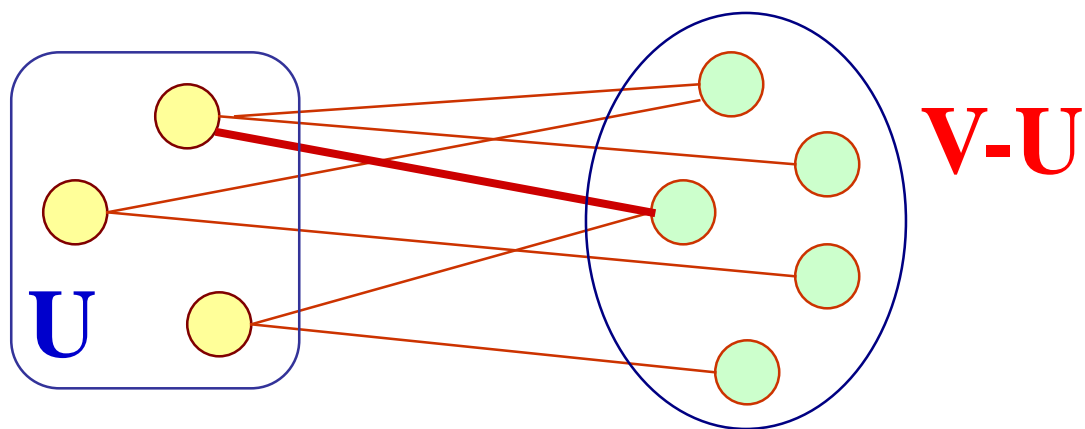
### ➤ MST性质的证明

- [反证]假设 $G$ 的任何一棵最小生成树都不包含 $(u,v)$ ，设 $T$ 是连通网上的一棵最小生成树，当将边 $(u,v)$ 加入到 $T$ 中时，由生成树的定义， $T$ 中必包含一条 $(u,v)$ 的回路。另一方面，由于 $T$ 是生成树，则在 $T$ 上必存在另一条边 $(u',v')$ ，且 $u$ 和 $u'$ 、 $v$ 和 $v'$ 之间均有路径相通。删去边 $(u',v')$ 便可消去上述回路，同时得到另一棵最小生成树 $T'$ 。但因为 $(u,v)$ 的代价不高于 $(u',v')$ ，则 $T'$ 的代价亦不高于 $T$ ， $T'$ 是包含 $(u,v)$ 的一棵最小生成树。





在生成树的构造过程中，图中 $n$ 个顶点分属两个集合：**已落在生成树上的顶点集 $U$** 和**尚未落在生成树上的顶点集 $V-U$** ，则应在所有连通 $U$ 中顶点和 $V-U$ 中顶点的边中选取权值最小的边。





## 4.4 最小生成树算法(cont.)

### ➤ 普里姆 (Prim) 算法

#### ■ 基本思想

- ① 首先从集合 $V$ 中任取一顶点(如顶点 $v_0$ )放入集合 $U$ 中。这时 $U=\{v_0\}$ , 边集 $TE=\{\}$
- ② 然后找出权值最小的边 $(u, v)$ , 且 $u \in U$ ,  $v \in (V-U)$ , 将边加入 $TE$ , 并将顶点 $v$ 加入集合 $U$
- ③ 重复上述操作直到 $U=V$ 为止。这时 $TE$ 中有 $n-1$ 条边,  $T=(U, TE)$ 就是 $G$ 的一棵最小生成树

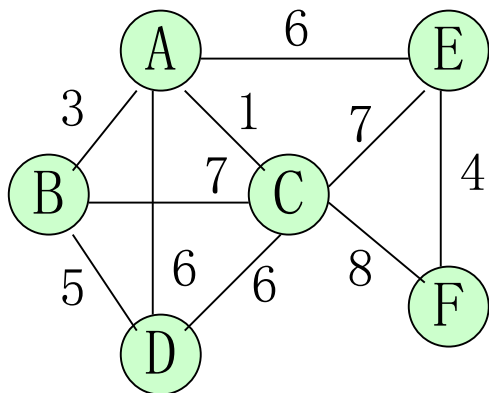
#### ■ 如何找到连接 $U$ 和 $V-U$ 的最短边

- 利用MST性质, 可以用下述方法构造候选最短边集: 对于 $V-U$ 中的每个顶点, 保存从该顶点到 $U$ 中的各顶点的最短边。





## Prim算法思想:



图G

黄色表示U的  
顶点，其他为  
V-U的顶点

A

V-U中各顶点到U的  
最短直接路径:

相邻顶点:

	A	B	C	D	E	F
A	0	3	1	6	6	$\infty$
B	3	0	7	5	$\infty$	$\infty$
C	1	7	0	6	7	8
D	6	5	6	0	$\infty$	$\infty$
E	6	$\infty$	7	$\infty$	0	4
F	$\infty$	$\infty$	8	$\infty$	4	0

初始化

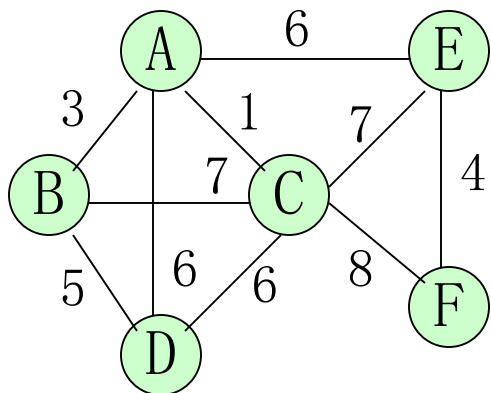
0	3	1	6	6	$\infty$
A	B	C	D	E	F

A	A	A	A	A	A
---	---	---	---	---	---



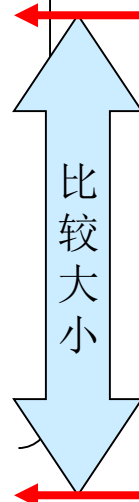


## Prim算法:



图G

	A	B	C	D	E	F
A	0	3	1	6	6	$\infty$
B	3	0	7	5	$\infty$	$\infty$
C	1	7	0	6	7	8
D	6	5	6	0	$\infty$	$\infty$
E	6	$\infty$	7	$\infty$	0	4
F	$\infty$	$\infty$	8	$\infty$	4	0



0	3	1	6	6	$\infty$
---	---	---	---	---	----------



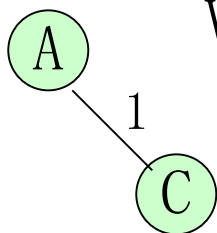
0	3	1	6	6	8
---	---	---	---	---	---

A B C D E F

A	A	A	A	A	A
---	---	---	---	---	---



A	A	A	A	A	C
---	---	---	---	---	---



V-U中各顶点到顶点集U的最短直接路径:

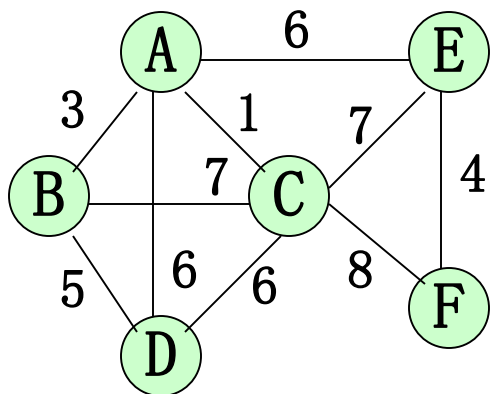
相邻顶点:



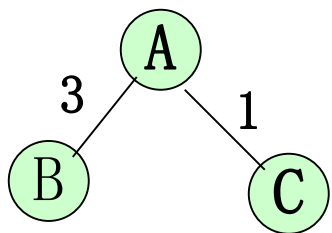




Prim算法:



图G



V-U中各顶点到顶点集U的最短直接路径

相邻顶点:

	A	B	C	D	E	F
A	0	3	1	6	6	$\infty$
B	3	0	7	5	$\infty$	$\infty$
C	1	7	0	6	7	8
D	6	5	6	0	$\infty$	$\infty$
E	6	$\infty$	7	$\infty$	0	4
F	$\infty$	$\infty$	8	$\infty$	4	0

比较大小

0	3	1	6	6	8
---	---	---	---	---	---

0	3	1	5	6	8
---	---	---	---	---	---

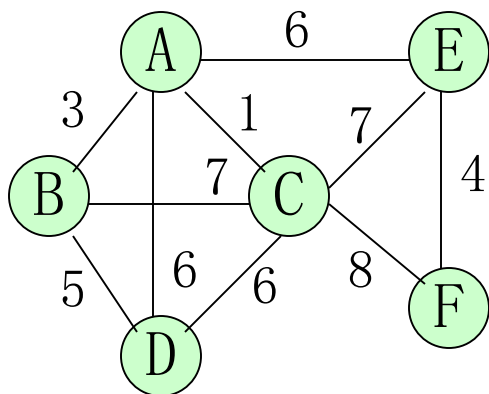
A	B	C	D	E	F
A	A	A	A	A	C

A	A	A	B	A	C
---	---	---	---	---	---

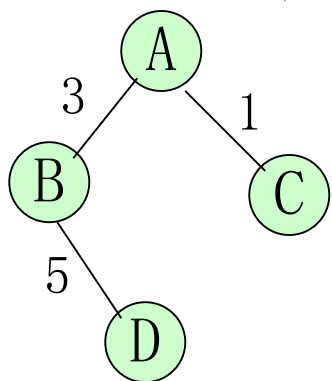




## Prime算法:



图G



V-U中各顶点到顶点  
集U的最短直接路径

相邻顶点:

	A	B	C	D	E	F
A	0	3	1	6	6	$\infty$
B	3	0	7	5	$\infty$	$\infty$
C	1	7	0	6	7	8
D	6	5	6	0	$\infty$	$\infty$
E	6	$\infty$	7	$\infty$	0	4
F	$\infty$	$\infty$	8	$\infty$	4	0

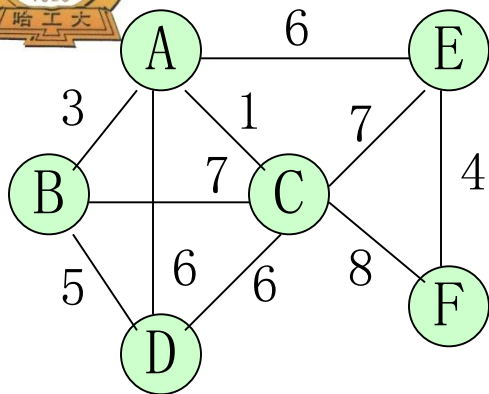
比较大小

0	3	1	5	6	8
0	3	1	5	6	8
A	B	C	D	E	F
A	A	A	B	A	C
A	A	A	B	A	C

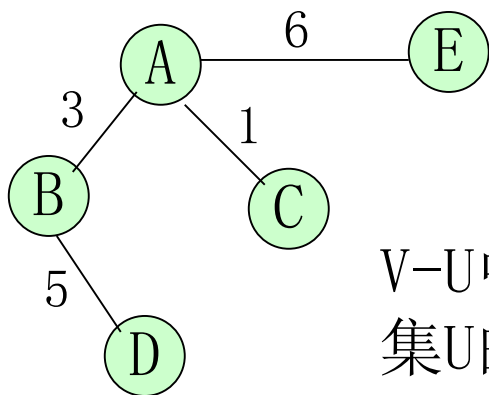




# Prime算法:



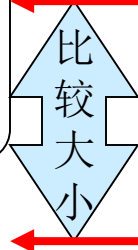
图G



V-U中各顶点到顶点集U的最短直接路径

相邻顶点:

	A	B	C	D	E	F
A	0	3	1	6	6	$\infty$
B	3	0	7	5	$\infty$	$\infty$
C	1	7	0	6	7	8
D	6	5	6	0	$\infty$	$\infty$
E	6	$\infty$	7	$\infty$	0	4
F	$\infty$	$\infty$	8	$\infty$	4	0

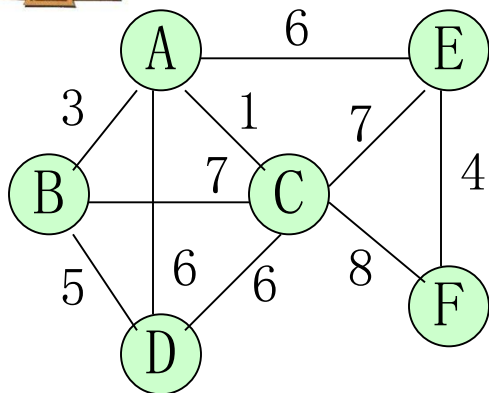


0	3	1	5	6	8
0	3	1	5	6	4
A	B	C	D	E	F
A	A	A	B	A	C
A	A	A	B	A	E

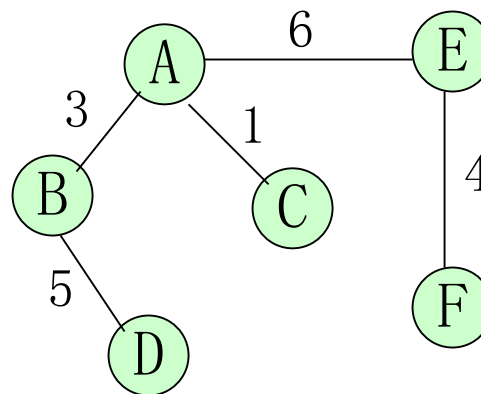




# Prime算法:



图G



U 和 V-U最短路径:

0	3	1	5	6	4
---	---	---	---	---	---



0	3	1	5	6	4
---	---	---	---	---	---

A B C D E F

A	A	A	B	A	E
---	---	---	---	---	---



A	A	A	B	A	E
---	---	---	---	---	---

相邻顶点:





## 4.4 最小生成树算法(cont.)

### ➡ 普里姆 (Prim) 算法的实现

#### ■ 数据结构

- 数组 **LOWCOST[n]**: 用来保存集合 **V-U** 中各顶点与集合 **U** 中顶点最短边的权值,  $\text{LOWCOST}[v] = \text{infinity}$  表示顶点  $v$  已加入最小生成树中;
- 数组 **CLOSEST[n]**: 用来保存依附于该边的 (集合 **V-U** 中各顶点与集合 **U** 中顶点的最短边) 在集合 **U** 中的顶点。

#### ■ 如何用数组 **LOWCOST[n]** 和 **CLOSEST[n]** 表示候选最短边集?

- $\text{LOWCOST}[i] = w$
  - $\text{CLOSEST}[i] = k$
- 表示顶点  $v_i$  和顶点  $v_k$  之间的权值为  $w$ , 其中:  $v_i \in V-U$  且  $v_k \in U$

#### ■ 如何更新?

$$\begin{cases} \text{LOWCOST}[j] = \min \{ \text{cost}(v_k, v_j) \mid v_j \in U, \text{LOWCOST}[j] \} \\ \text{CLOSEST}[j] = k \end{cases}$$





## 4.4 最小生成树算法(cont.)

### ■ 实现步骤:

1. 初始化两个辅助数组LOWCOST和CLOSEST;
2. 输出顶点 $v_0$ , 将顶点 $v_0$ 加入集合U中;
3. 重复执行下列操作 $n-1$ 次
  - 3.1 在LOWCOST中选取最短边, 取CLOSEST中对应的顶点序号 $k$ ;
  - 3.2 输出顶点 $k$ 和对应的权值;
  - 3.3 将顶点 $k$ 加入集合U中;
  - 3.4 调整数组LOWCOST和CLOSEST;

$$\begin{cases} \text{LOWCOST}[j] = \min \{ \text{cost}(v_k, v_j) \mid v_j \in U, \text{LOWCOST}[j] \} \\ \text{CLOSEST}[j] = k \end{cases}$$





## 4.4 最小生成树算法(cont.)

### ➡ 普里姆 (Prim) 算法的实现

```


void Prim(Costtype C[n+1][n+1] )
{ costtype LOWCOST[n+1]; int CLOSEST[n+1]; int i,j,k; costtype min;
  for( i=2; i<=n; i++ )
  {   LOWCOST[i] = C[1][i];   CLOSEST[i] = 1;   }
  for( i = 2; i <= n; i++ )
  {   min = LOWCOST[i];
      k = i;
      for( j = 2; j <= n; j++ )
          if ( LOWCOST[j] < min )
              { min = LOWCOST[j];   k=j; }
      cout << "(" << k << "," << CLOSEST[k] << ")" << endl;
      LOWCOST[k] = infinity ;
      for ( j = 2; j <= n; j++ )
          if ( C[k][j] < LOWCOST[j] && LOWCOST[j] < infinity )
              {   LOWCOST[j]=C[k][j];   CLOSEST[j]=k;   }
  }
}
/* 时间复杂度:  $O(|V|^2)$ 

```



## 算法分析

分析Prim算法，该算法由两个并列的循环组成，第一个循环次数为vex\_num(即顶点的个数 $n$ )；第二个循环，外层循环的次数为 $n-1$ ，内层的循环次数为 $n$ 。所以总体来说，Prim的时间复杂度为 $O(n^2)$ ，并且该算法与图中边数的多少无关，所以该算法适合于求边稠密的图的最小生成树。







## 4.4 最小生成树算法(cont.)

### ➤ 克鲁斯卡尔 (Kruskal) 算法

#### ■ 基本思想:

- 设无向连通网为 $G=(V, E)$ , 令 $G$ 的最小生成树为 $T=(U, TE)$ , 其初态为 $U=V, TE=\{ \}$ ,
- 然后, 按照边的权值由小到大的顺序, 依次考察 $G$ 的边集 $E$ 中的各条边。
- 若被考察的边连接的是两个不同连通分量, 则将此边作为最小生成树的边加入到 $T$ 中, 同时把两个连通分量连接为一个连通分量;
- 若被考察的边连接的是同一个连通分量, 则舍去此边, 以免造成回路,
- 如此下去, 当 $T$ 中的连通分量个数为1时, 此连通分量便为 $G$ 的一棵最小生成树。





## 4.4 最小生成树算法(cont.)

### ➤ 克鲁斯卡尔 (Kruskal) 算法

#### ■ 实现步骤:

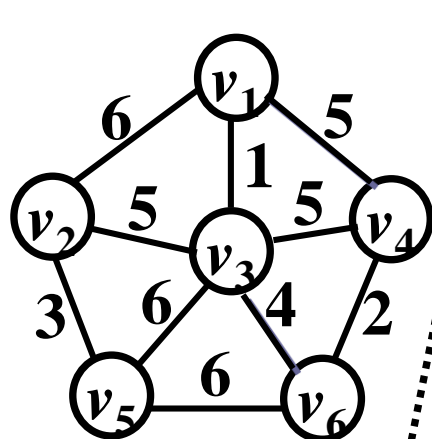
1. 初始化:  $U=V$ ;  $TE=\{ \}$ ;
2. 循环直到T中的连通分量个数为1
  - 2.1 在E中选择最短边( $u, v$ );
  - 2.2 如果顶点 $u$ 、 $v$ 位于T的两个不同连通分量, 则
    - 2.2.1 将边( $u, v$ )并入TE;
    - 2.2.2 将这两个连通分量合为一个;
  - 2.3 在E中标记边( $u, v$ ), 使得( $u, v$ )不参加后续最短边的选取



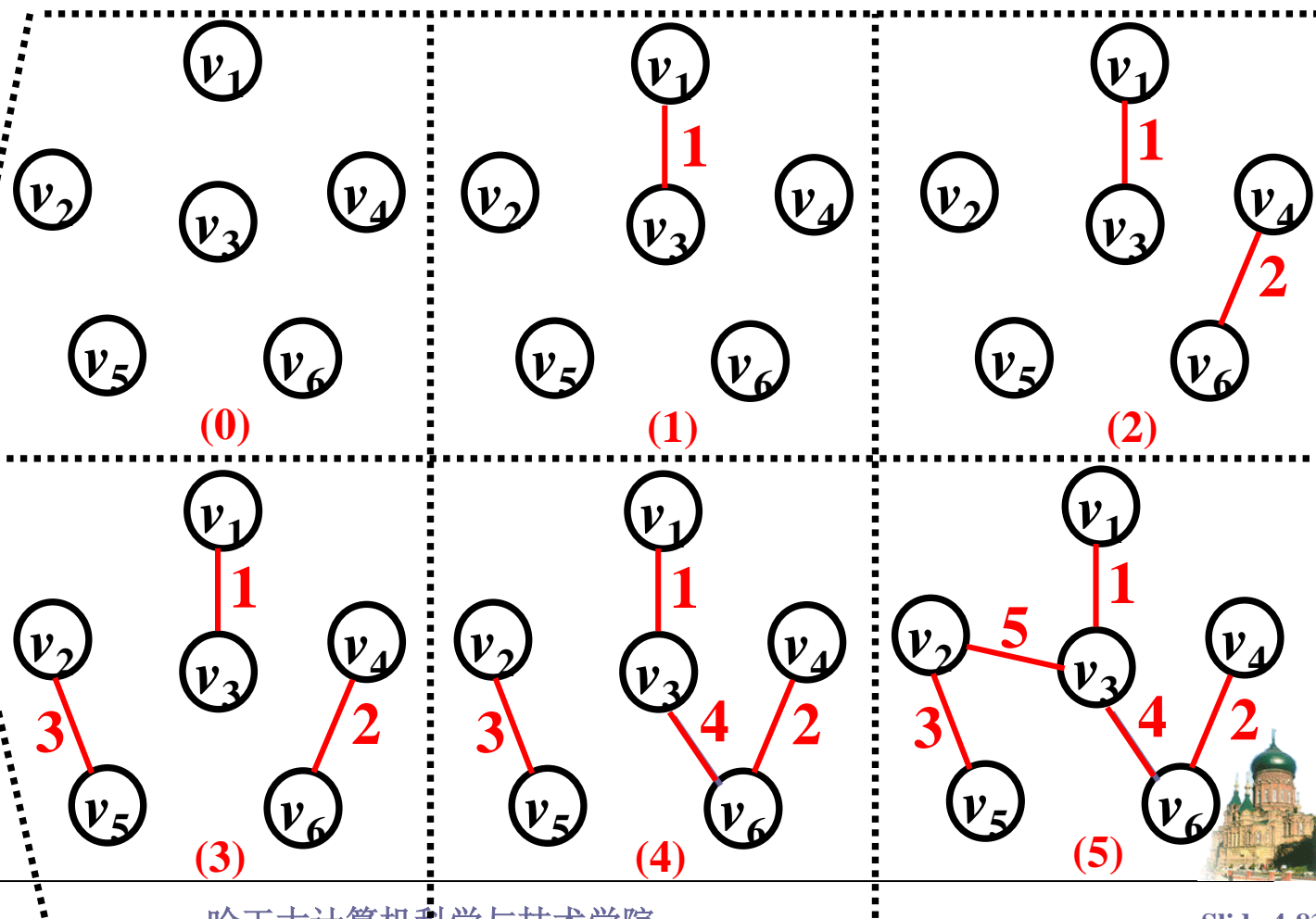


## 4.4 最小生成树算法(cont.)

$(v_1, v_3)$	$(v_4, v_6)$	$(v_2, v_5)$	$(v_3, v_6)$	$(v_1, v_4)$	$(v_3, v_4)$	$(v_2, v_3)$	$(v_1, v_2)$	$(v_3, v_5)$	$(v_5, v_6)$
1	2	3	4	5	5	5	6	6	6



$O(|E| \cdot \log |E|)$



## 算法分析

Kruscal算法至多对 $e$ 条边各扫描一次，假若用堆来存放网中的边，则每次选择最小代价的边仅需 $O(\log e)$ 的时间。又生成树 $T$ 的每个连通分量可看成是一个等价类，则构造 $T$ 加入新的边的过程类似于求等价类的过程，有算法保证其时间复杂度可以达到 $O(e \log e)$ 。因此，Kruscal算法的时间复杂度为 $O(e \log e)$ 。所以该算法适合求边稀疏的图的最小生成树。





## kraskal算法(要求边权值以不减顺序排列)

```
typedef struct edge{
    int bgn , end, wet;
}Egde;
```

```
void Kraskal (Egde edges[], int e)
{
    int father[], bnf, edf, i;
    for (i=1;i<=e;i++)
        father[i]=0;

    for ( i=1;i<=e; i++)
    {
        bnf= Find (father, eds[i]. bgn);
        edf= Find (father, edges[i]. end);
        if (bnf != edf)
            father [bnf] = edf;
    }
}
```

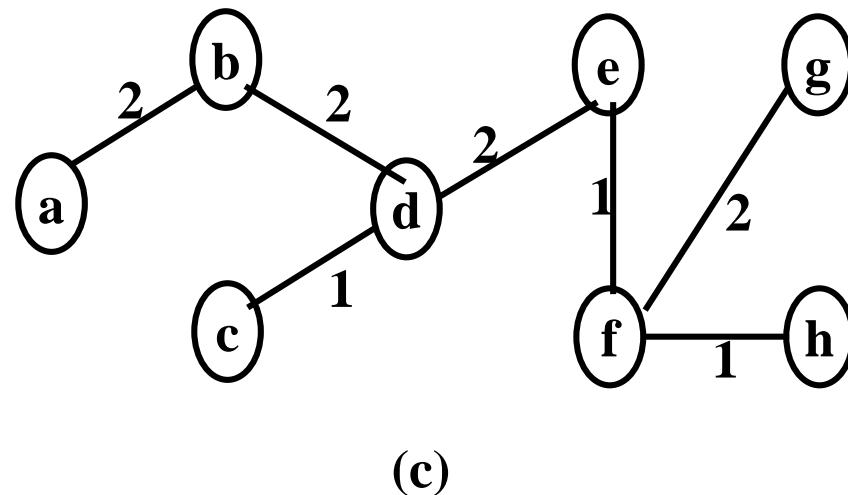
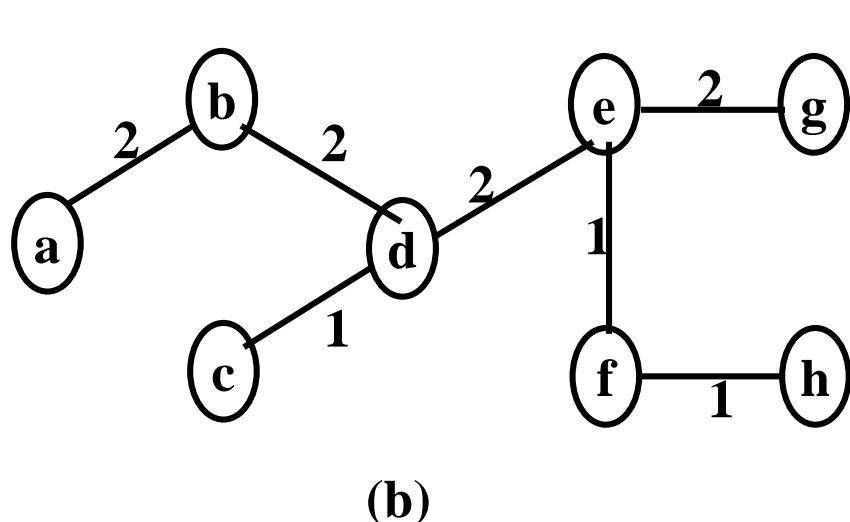
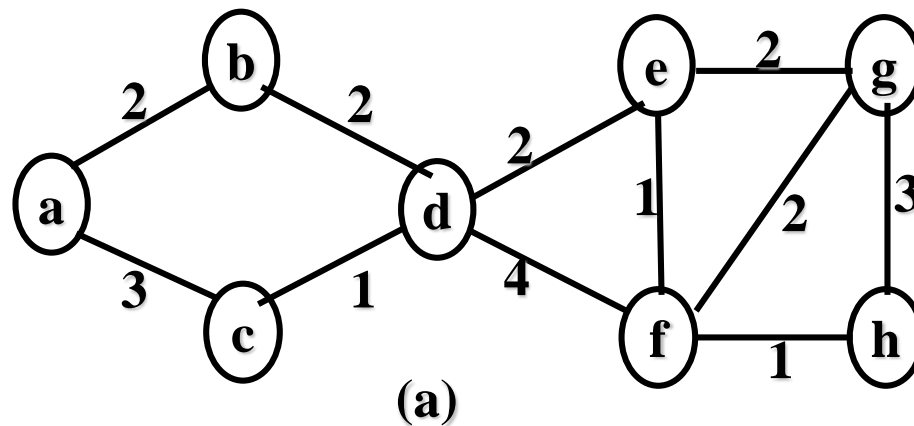
```
void main()
{
    Edge edges [max];
    e= Getedges (edges); //边的个数
    Sort( edges, e);
    Kruskal (edges, e);
}
```

```
int Find( int father[], int v)
{
    int f = v;
    while ( father[f] >0)
        f = father [f];
    return (f);
}
```





例



当各边有相同权值时，由于选择的任意性，产生的生成树可能不唯一  
当各边的权值不相同时，产生的生成树是唯一的。





### 作业一：农夫过河

农夫需要把狼、羊、菜和自己运到河对岸去，只有农夫能够划船，而且船比较小，除农夫之外每次只能运一种东西，还有一个棘手问题，就是如果没有农夫看着，羊会偷吃菜，狼会吃羊。请考虑一种方法，让农夫能够安全地安排这些东西和他自己过河。

### 作业二：迷宫问题

给定一个迷宫，入口为左上角，出口为右下角，问是否有路径从入口到出口，若有则输出一条这样的路径，并判断是否为最佳路径。注意移动可以从上、下、左、右、上左、上右、下左、下右八个方向进行。迷宫输入0表示可走，输入1表示墙。可以用1将迷宫围起来避免边界问题。



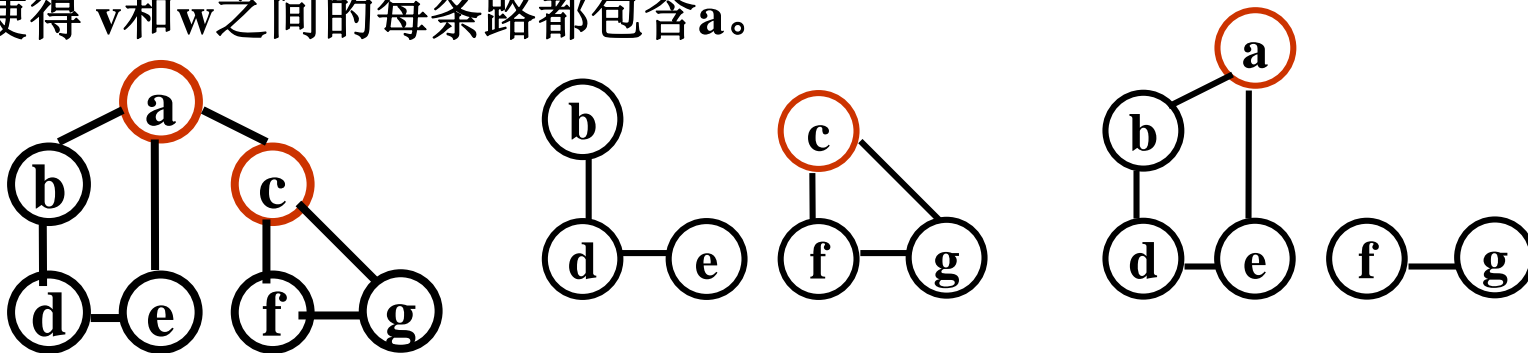


## 4.5 无向图的双连通性

### 4.5.1 无向图的双连通分量 (Biconnected Component)

设  $G = (V, E)$  是一个连通的无向图

称顶点  $a \in V$  是关节点 (articulation point), 如果存在  $v, w \in V$ ,  $v \neq w \neq a$  且使得  $v$  和  $w$  之间的每条路都包含  $a$ 。



若在删去顶点  $v$  以及与之相邻的边之后, 将图的一个连通分量分割成两个或两个以上的连通分量, 则称该顶点为关节点。

**定义** 若对  $V$  中每个不同的三元组  $v, w, a$ ; 在  $v$  和  $w$  之间都存在一条不包含  $a$  的路, 就说  $G$  是双连通的 (biconnected)

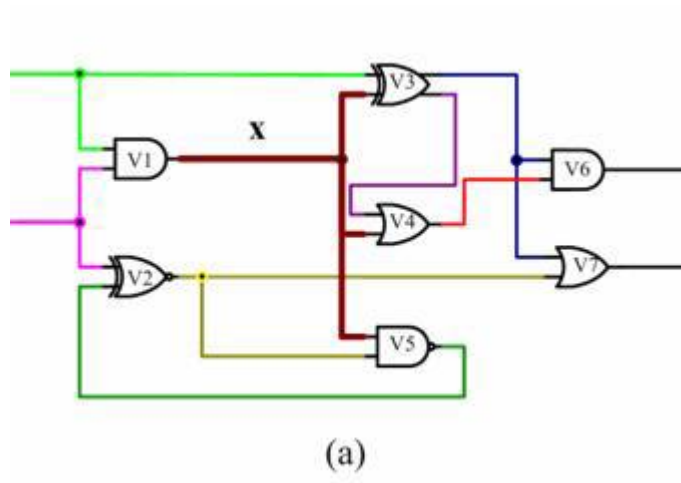
没有关节点的连通图称为双连通图。



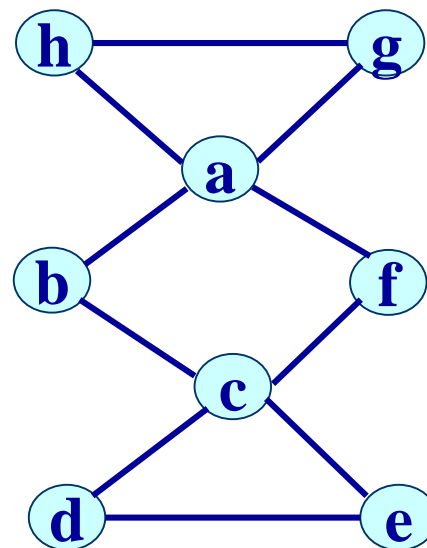




- 双连通的无向图是连通的，但连通的无向图未必双连通。
- 一个连通的无向图是双连通的，当且仅当它没有关节点。
- 在双连通图上，任何一对顶点之间至少存在有两条路径，在删去某个顶点及与该顶点相关联的边时，也不破坏图的连通性。
- 一个连通图G如果是重连通图，那么它可以包括几个双连通分量。



下列连通图中，



顶点a和顶点c是关节点

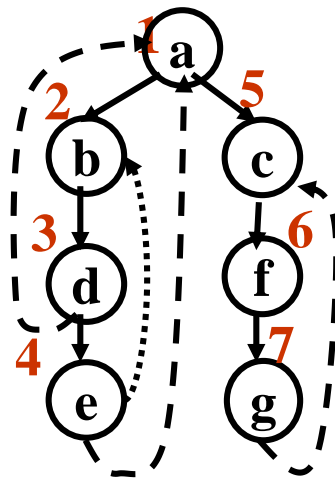
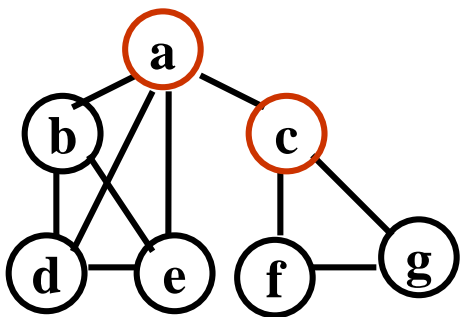




## 4.5.2 求关节点—对图进行一次先深搜索便可求出所有的关节点

由先深生成树可得出两类关节点的特性：

- 若生成树的根有两株或两株以上子树，则此根结点必为关节（第一类关节点）。因为图中不存在连接不同子树中顶点的边，因此，若删去根顶点，生成树变成生成森林。
- 若生成树中非叶顶点 $v$ ，其某株子树的根和子树中的其它结点均没有指向 $v$ 的祖先的回退边，则 $v$ 是关节点（第二类关节点）。因为删去 $v$ ，则其子树和图的其它部分被分割开来



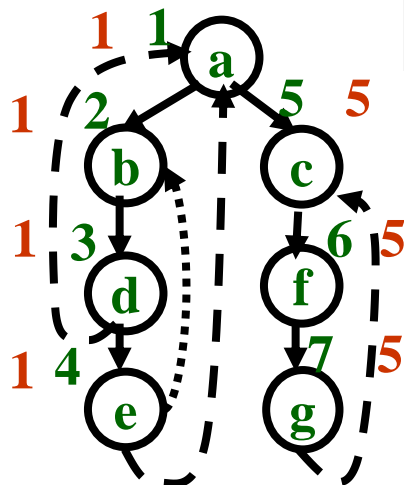
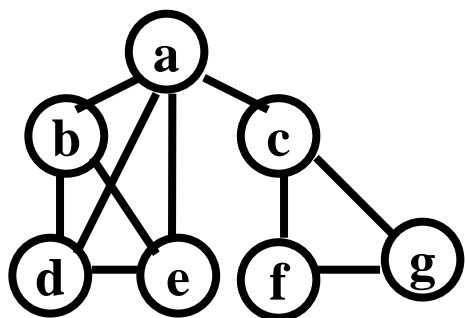


定义  $\text{low}[v]$  设对连通图  $G=(V,E)$  进行先深搜索的先深编号为  $\text{dnf}[v]$ , 产生的先深生成树为  $S=(V,T)$ ,  $B$  为回退边之集。对每个顶点  $v$ ,  $\text{low}[v]$  定义如下:

$$\text{low}[v] = \min \left\{ \begin{array}{l} \text{dnf}[v], \text{dnf}[w], \text{low}[y] \end{array} \right.$$

$(w, v) \in B$ ,  $w$  是顶点  $v$  在先深生成树上有回退边连接的祖先结点;

$(v, y) \in T$ ,  $y$  是顶点  $v$  在先深生成树上的孩子顶点



$\text{dnf}[ ]$   $\text{low}[ ]$

序	结点	$\text{dnf}[v]$	$\text{dnf}[w]$	$\text{low}[y]$	$\min \{ \}$
1	e	4	1, 2		1
2	d	3	1	1	1
3	b	2		1	1
4	g	7	5		5
5	f	6		5	5
6	c	5		5	5
7	a	1		1, 5	1





## 算法4.5 求无向图的双连通分量

**输入：**连通的无向图 $G=(V, E)$ 。 $L[v]$ 表示关于 $v$ 的邻接表

**输出：** $G$ 的所有双连通分量，每个连通分量由一序列的边组成。

**算法要点：**

- 1.计算先深编号：对图进行先深搜索，计算每个结点 $v$ 的先深编号 $dnf[v]$ ，形成先深生成树 $S=(V, T)$ 。
- 2.计算 $low[v]$ ：在先深生成树上按后根顺序进行计算每个顶点 $v$ 的 $low[v]$ ， $low[v]$ 取下述三个结点中的最小者：
  - (1)  $dnf[v]$ ;
  - (2)  $dnf[w]$ ，凡是有回退边 $(v, w)$ 的任何结点 $w$ ;
  - (3)  $low[y]$ ，对 $v$ 的任何儿子 $y$ 。
- 3.求关节点：
  - (1)树根是关节点，当且仅当它有两个或两个以上的儿子（第一类关节点）；
  - (2)非树根结点 $v$ 是关节点当且仅当 $v$ 有某个儿子 $y$ ，使 $low[y] \geq dnf[v]$ （第二类关节点）。





### 求双连通分量的算法

**void main()**

```
{    count = 1;
    for (all  $v \in V$ ) mark  $v$  “new”;
    searchB(  $v_0$  );
```

// $v_0$ 是 $V$ 中任一结点,建立以 $v_0$ 为根的先深生成树,计算low( $v$ )

**void searchB( $v$ )**

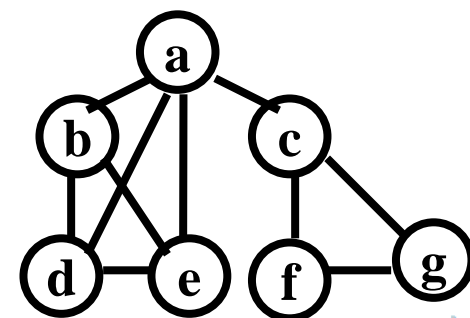
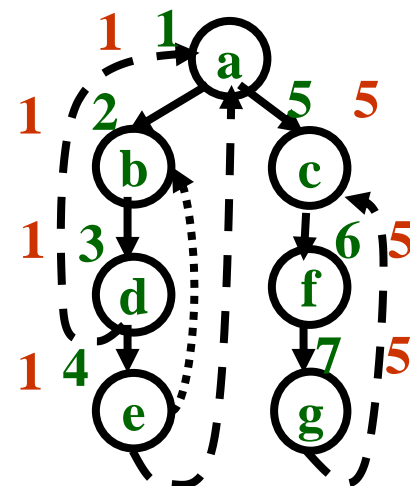
```
{
1    mark  $v$  “old”;
2    dfn[ $v$ ] = count; count++;
3    low[ $v$ ] = dfn[ $v$ ];
4    for (each  $w \in L[v]$ )
```





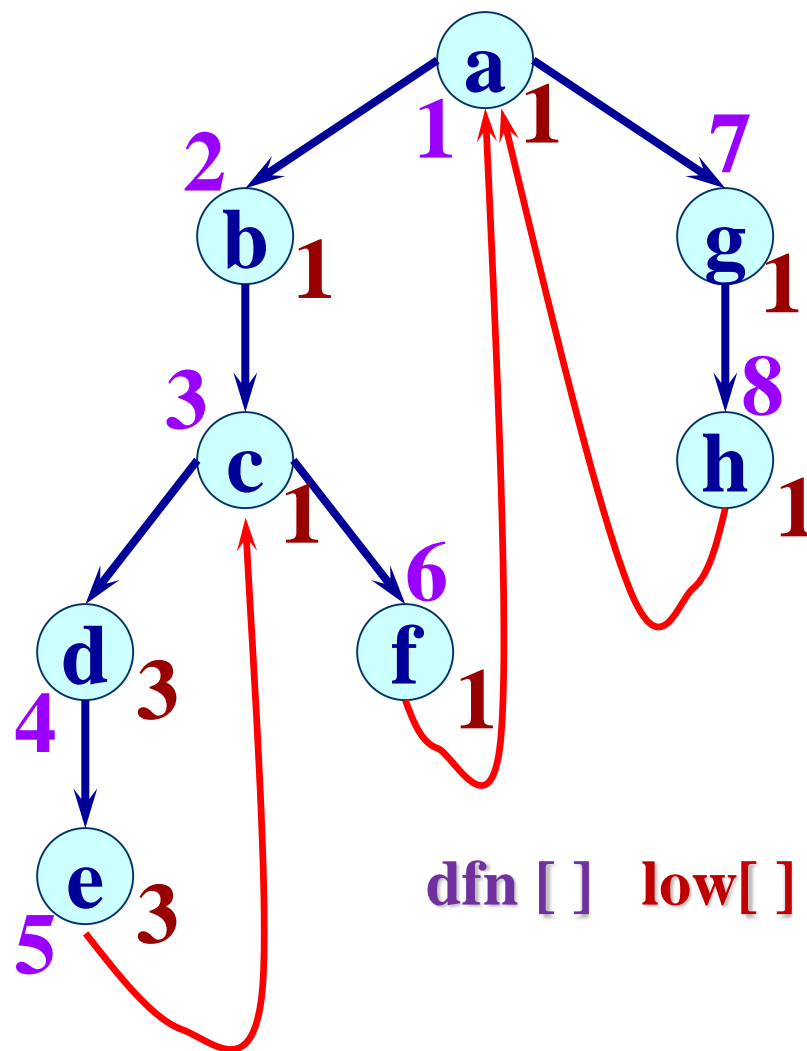
```

7  if (w is marked “new”)
    {
8      father [w]=v;
9      searchB (w);
10     low [v]= min { low [v], low [w] };
11     if (low[w]>=dfn[v]) //表明w及子孙均无指向v的祖先的回
        边，v是关节点
        cout<<“a biconnected component”;
    }
12 else
    if (w != father [v]) //(v ,w)是回退边
13     low [v]= min { low [v], dfn [w] };
    }
    
```





# 深度优先生成树



dfn [ ] low [ ]





有一群牛，总数为 $N$  ( $N \leq 10000$ )，题目数据给出牛之间的关系，比如说1仰慕2，2仰慕3等等，设这种仰慕是可以传递的，如果1仰慕2，那么1也会同时仰慕2仰慕的那些牛，如果一头牛被所有的牛都仰慕，那么它将是受欢迎的牛，题目要求的是有多少牛是“最受欢迎的”。

这道题目大家第一眼看到可能感觉直接模拟，但是由于数据量巨大，模拟的话肯定是过不了的，而且题目中还会出现环路的情况，比如 $1 \Rightarrow 2, 2 \Rightarrow 3, 3 \Rightarrow 1$ ，所以这解这道题最好的方法是使用有向图的强连通分量。

在同一个强连通分量里的所有的牛之间是互相仰慕的，我们将其缩为一点，并且只记录这一点的牛的数量，如果有最受欢迎的牛存在，那么这个图将是连通图，并且出度为零的点只有一个，我们可以用并查集来判是否连通，然后计算每个节点的出度，即可求出最受欢迎的牛的数量。

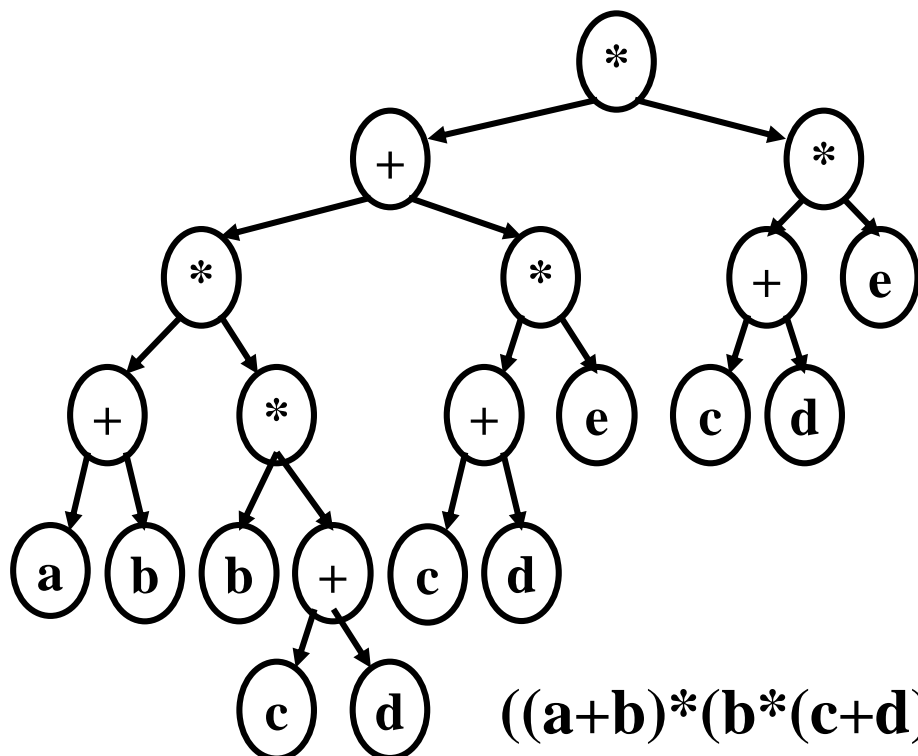






## 4.6 拓扑(topology)排序算法

- **无环路有向图**：不存在环路的有向图的简称。
- **注意**：无环路的有向图对应的无向图可能存在环路。
- 无环路的有向图可以描述**含有公共子式的表达式**(节省空间)。
- 无环路的有向图可用于表示**偏序集**。



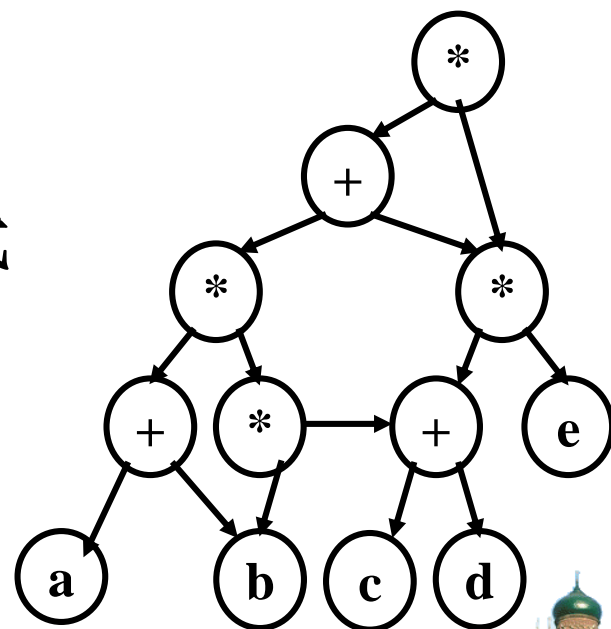
公共子式

b

(c+d)

(c+d)\*e

$((a+b)*(b*(c+d))+(c+d)*e)*((c+d)*e)$





## 4.6 拓扑排序算法(cont.)

➡ **偏序关系**：若集合 $X$ 上的关系 $R$ 是**自反的**、**反对称的**和**传递的**

- 自反性：任意 $x \in X$ ,  $(x, x) \in R$

- 反对称性：任意 $x, y \in X$ , 若 $(x, y) \in R$ 且 $(y, x) \in R$ , 则 $x=y$

- 传递性：任意 $x, y, z \in X$ ,  $(x, y) \in R$ 且 $(y, z) \in R$ , 则 $(x, z) \in R$

则称 $R$ 是集合 $X$ 上的**偏序关系**。

➡ **全序关系**：

- 设 $R$ 是集合 $X$ 上的偏序关系，如果对每个 $x, y \in X$ , 必有 $(x, y) \in R$ 或 $(y, x) \in R$ , 则称 $R$ 是集合 $X$ 上的**全序关系**

➡ **直观上**, **偏序**指集合上只有**部分元素**之间可比较, 而**全序**是指**全体元素**均可比较。





## 4.6 拓扑排序算法(cont.)

### ➤ 如何用无环路的有向图表示偏序关系？

- 设 $R$ 是有穷集合 $X$ 上的偏序关系，对 $X$ 中每个 $v$ ，用一个以 $v$ 为标号的顶点表示，由此构成顶点集 $V$ ；对任意 $(u, v) \in R$ ，( $u \neq v$ )，由对应两个顶点建立一条有向边，由此构成边集 $E$ ，则 $G = (V, E)$ 是无环路有向图。

### ➤ 拓扑排序：是由某个集合上的一个偏序得到该集合上的一个全序的过程。所得到的线性序列称为拓扑序列。

### ➤ AOV网：在一个表示工程的有向图中，用顶点表示活动，用弧表示活动之间的优先关系，称这样的有向图为顶点表示活动的网，简称AOV网。

- AOV网中的弧表示活动之间存在的某种制约关系。
- AOV网中不能出现回路。
- 在AOV网中，若从顶点 $i$ 到 $j$ 有一条有向路，则称 $i$ 为 $j$ 的前驱， $j$ 为 $i$ 的后继。若 $(i, j) \in E$ ，则 $i$ 称为 $j$ 的直接前驱， $j$ 称为 $i$ 的直接后继。



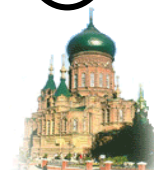
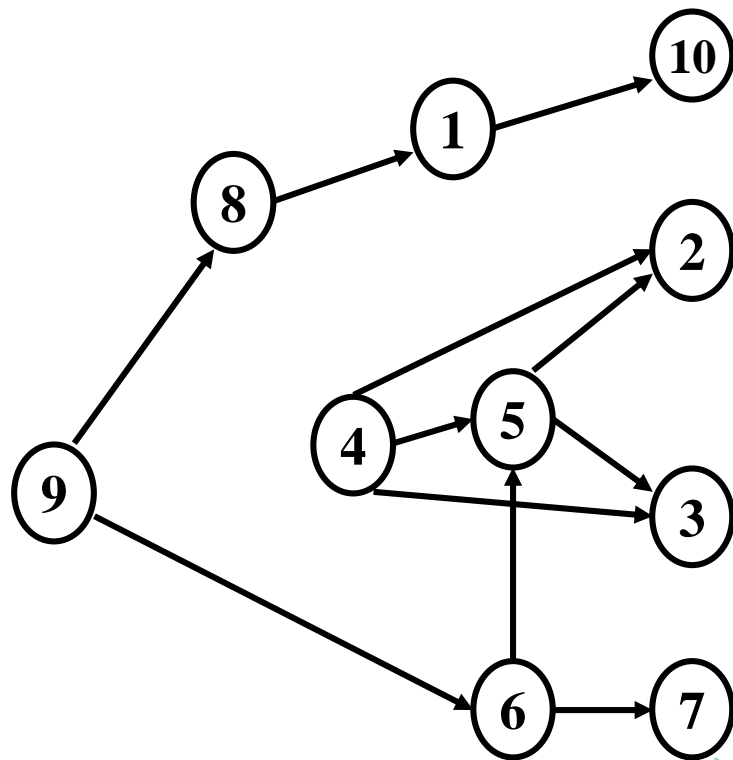


## 4.6 拓扑排序算法(cont.)

### ➡ AOV网示例:

- 课程及课程间的先修关系是偏序关系，可以用AOV网表示

课程代号	课程名称	先修课代号
1	计算机原理	8
2	编译原理	4,5
3	操作系统	4,5
4	程序设计	无
5	数据结构	4,6
6	离散数学	9
7	形式语言	6
8	电路基础	9
9	高等数学	无
10	计算机网络	1

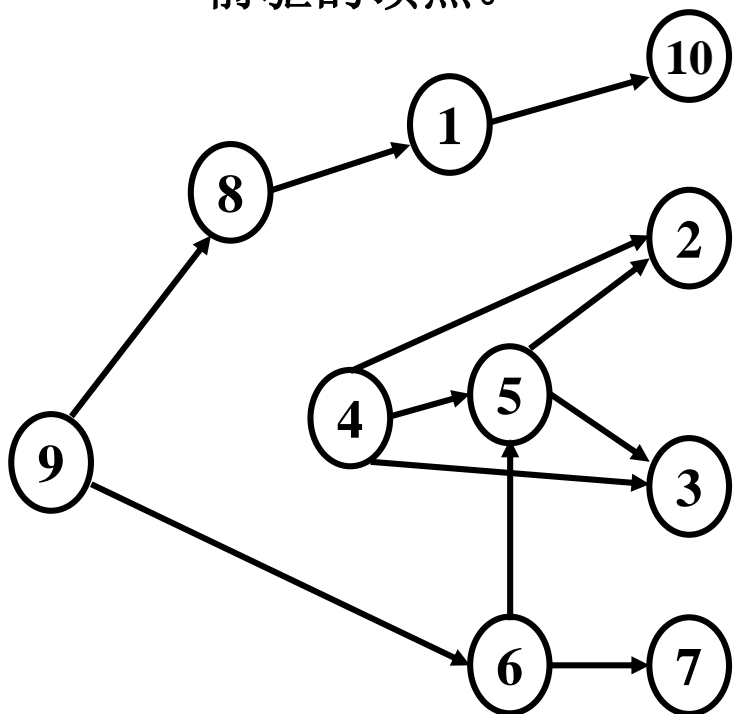




## 4.6 拓扑排序算法(cont.)

➔ 利用**AOV**网进行拓扑排序的基本思想：

- (1) 从**AOV**网中选择一个没有前驱的顶点并且输出它；
- (2) 从**AOV**网中删去该顶点和所有以该顶点为尾的弧；
- (3) 重复上述两步，直到全部顶点都被输出，或**AOV**网中不存在没有前驱的顶点。



任何无环路的**AOV**网，其顶点都可以排成一个拓扑序列，并且其拓扑序列不一定是唯一的。

(9 8 6 1 4 5 7 2 3 10)

(9 4 8 6 5 1 10 2 3 7)





## 4.6 拓扑排序算法(cont.)

➔ **拓扑排序算法**——实质是广度优先搜索算法

■ **输入**：有向图的邻接表， **输出**：所有顶点组成的拓扑序列

■ **算法实现步骤**：（使用队列）

1. 建立入度为零的顶点排队
2. 扫描顶点表，将入度为0的顶点入队；
3. while（排队不空）{  
    输出队头结点；  
    记下输出结点的数目；  
    删去与之关联的出边；  
    若有入度为0的结点，入队  
}
4. 若输出结点个数小于 $n$ ，则输出有环路；否则拓扑排序正常结束。

☞ 图中还有未输出的顶点，但已跳出循环处理。说明图中还剩下一些顶点，它们都有直接前驱。这时网络中必存在有向环；或

☞ 全部顶点均已输出，拓扑有序序列形成，拓扑排序完成。





## 4.6 拓扑排序算法(cont.)

➡ **拓扑排序算法**——实质是**广度优先搜索算法**

```
void Topologicalsort( AdjGraph G )
```

```
{ Queue Q ; nodes = 0 ;
```

```
  MakeNull( Q ) ;
```

```
  for( v=1; v<=G.n ; ++v )
```

```
    if ( indegree[v] ==0 ) EnQueue( v, Q ) ;
```

```
  while ( !Empty( Q ) ) {
```

```
    v = Front(Q) ;
```

```
    DeQueue( Q ) ;
```

```
    cout << v ; nodes ++ ;
```

```
    for( 邻接于 v 的每个顶点 w )
```

```
      if( !(--indegree[w])) EnQueue(w,Q) ;
```

```
  }
```

```
  if ( nodes < n ) cout<<“图中有环路” ;
```

```
}
```

```
for(p=G.verlist[v].firstedge; p; p=p->next)
{ w=p->adjvex;
  if( !(--indegree[w]))
    EnQueue(w,Q) ;
}
```





## 4.6 拓扑排序算法(cont.)

### ➤ 关于广度优先拓扑排序的几点说明

#### ■ 与先广搜索的差别：

- 搜索起点是入度为0的顶点；
- 需判断是否有环路；
- 需对访问并输出的顶点计数（引入计数器**nodes**）。
- 需删除邻接于 **v** 的边（引入数组**indegree[ ]**或在顶点表中增加一个属性域**indegree**）。

#### ■ 也可以采用**栈**数据结构进行广度优先拓扑排序。

#### ■ 亦可采用**无后继顶点优先**的拓扑排序算法

#### ■ 也可以利用**DFS遍历**进行拓扑排序







## 4.6 拓扑排序算法(cont.)

### ➤ 利用栈结构进行拓扑排序

■ **输入**：有向图的邻接表， **输出**：所有顶点组成的拓扑序列

■ **算法实现步骤**：（使用栈）

1. 建立入度为零的顶点**栈**
2. 扫描顶点表，将入度为0的顶点**栈**；
3. while（**栈**不空）{  
    输出栈顶结点；  
    记下输出结点的数目；  
    删去与之关联的出边；  
    若有入度为0的结点，入**栈**  
}
4. 若输出结点个数小于 $n$ ，则输出有环路；否则拓扑排序正常结束。





## 4.6 拓扑排序算法(cont.)

➡ 利用栈结构进行拓扑排序

```
void Topologicalsort( AdjGraph G )
{   MakeNull( S ) ; count = 0 ;
    for( v=0; v<n ; ++v )
        if ( !indegree[v] ) Push( v, S ) ;
    while ( !Empty( S ) ) {
        v = Pop ( S ) ; printf( v ); ++count ;
        for( 邻接于 v 的每个顶点 w ) {
            if( !(--indegree[w]))
                Push(S, w) ;
        }
    }
    if ( count < n ) cout<<“图中有环路” ;
}
```





## 4.6 拓扑排序算法(cont.)

### ➡ 基于DFS的拓扑排序

```
void topodfs ( v )  
{ Push( v ,S );  
  mark[v]=True;  
  for ( L[v] 中的每一个顶点w)  
    if ( mark[w] = False )  
      topodfs ( w );  
  printf ( Top( S ) );  
  Pop ( S );  
}
```

```
void dfs-topo ( GRAPH L )  
{ MakeNull( S );  
  for( u=1;u<=n; u++)  
    mark[u]=False;  
  for( u=1;u<=n;u++)  
    if ( !mark[u] )  
      topodfs( u );  
}
```

**思想：**借助栈，在DFS中，把第一次遇到的顶点入栈，到达某一顶点递归返回时，从栈中弹出顶点并输出。





## 4.7 关键路径算法

### 案例场景

- 某软件公司承接一家企业的信息系统集成业务。经过公司董事会的讨论，决定任命你作为该系统集成项目的项目经理，你接到任务后，应该制定进度表，这样项目才可以依照进度表继续下去。
- 在与项目团队成员探讨后，假设已经确认了共9件事，11项基本活动。所有这些活动的名称、完成每项活动所需的时间，以及与其他活动之间的约束关系如右侧表：

活动名称	必需的时间(天)	前置任务
a1	6	
a2	4	
a3	5	
a4	1	a1
a5	1	a2
a6	2	a6
a7	9	a4,a5
a8	7	a4,a5
a9	4	a6
a10	2	a7
a11	4	a8,a9

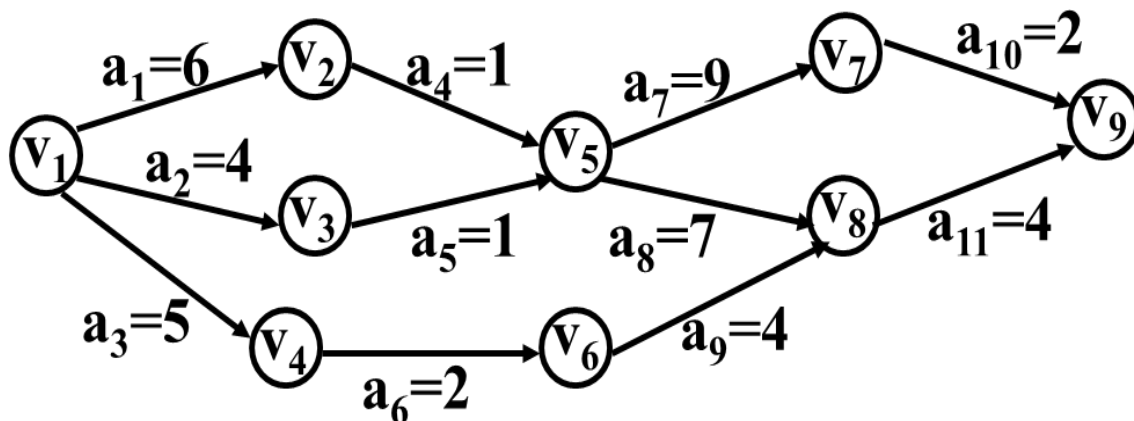




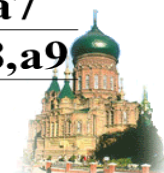
## 4.7 关键路径算法(cont.)

- ➡ **问题1:** 如何描述项目进度?
- ➡ **问题2:** 完成整个项目至少需要多少时间?
- ➡ **问题3:** 如果在任务a5上推迟了3天, 对项目进度有何影响?

作为项目经理, 你将如何处理这个问题?



活动名称	必需的时间(天)	前置任务
a1	6	
a2	4	
a3	5	
a4	1	a1
a5	1	a2
a6	2	a3
a7	9	a4,a5
a8	7	a4,a5
a9	4	a6
a10	2	a7
a11	4	a8,a9

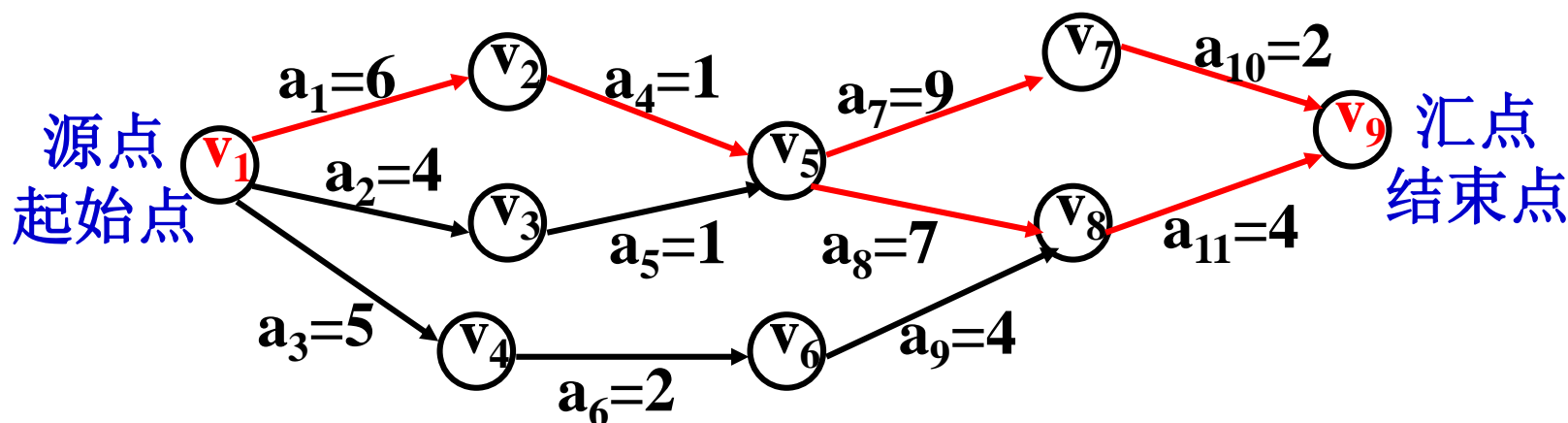




## 4.7 关键路径算法(cont.)

### ➡ AOE网 (Activity On Edge Network)

- 在带权的有向图中，用顶点表示**事件**，用边表示**活动**，边上权表示活动的**开销**（如**持续时间**），则称此有向图为**边表示活动的网络**，简称**AOE网**。
- 下图是有**11项** 活动，**9个**事件的**AOE网**，每个事件表示在它之前的活动已经完成, 在它之后的活动可以开始。

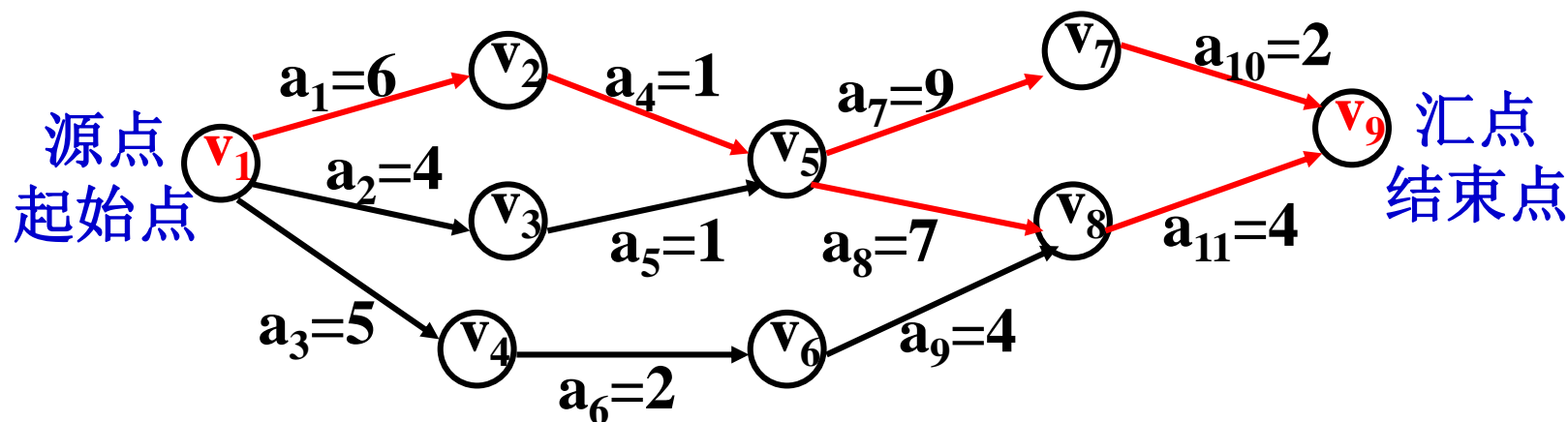




## 4.7 关键路径算法(cont.)

### ➔ AOE网的性质

- 只有在某个顶点所代表的事件发生后，从该顶点出发的各有向边代表的活动才能开始；
- 只有在进入某一顶点的各有向边代表的活动已经结束，该顶点所代表的事件才能发生；
- 表示实际工程计划的AOE网应该是**无环的**，并且存在唯一的入度为0的开始顶点（**源点**）和唯一的出度为0的结束点（**汇点**）。



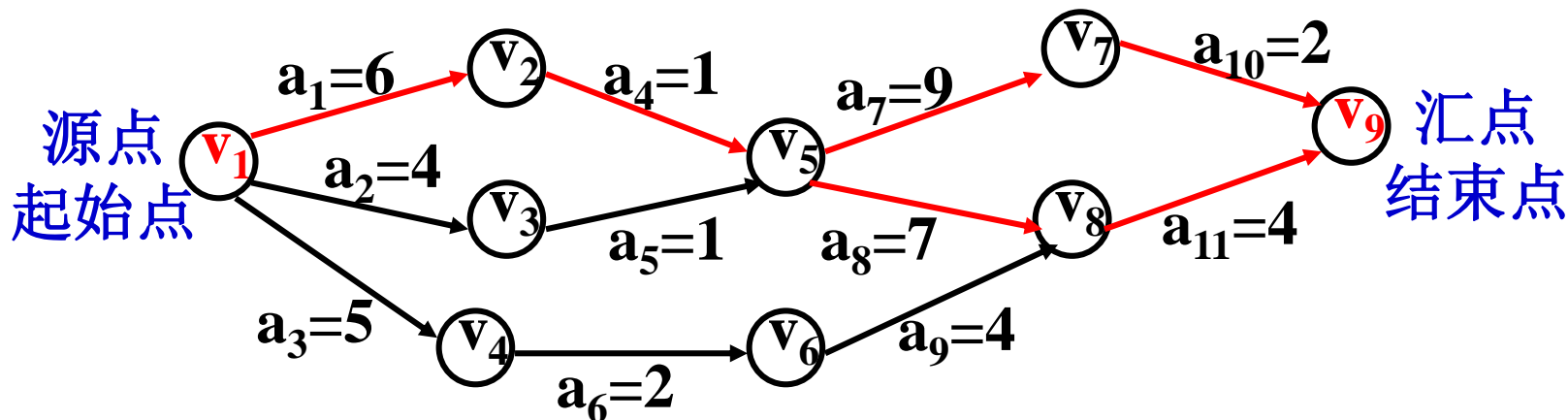


## 4.7 关键路径算法(cont.)

### ➡ AOE网研究的主要问题:

- 如果用AOE网表示一项工程，那么仅仅考虑工程之间的**优先关系**还不够，更多地是关心整个工程完成的最短时间是多少，哪些**活动的延迟将影响整个工程进度**，而**加速这些活动能否提高整个工程的效率**，因此AOE网有待研究的问题是：

- (1)完成整个工程至少需要多少时间？
- (2)哪些活动是影响工程进度的关键活动？



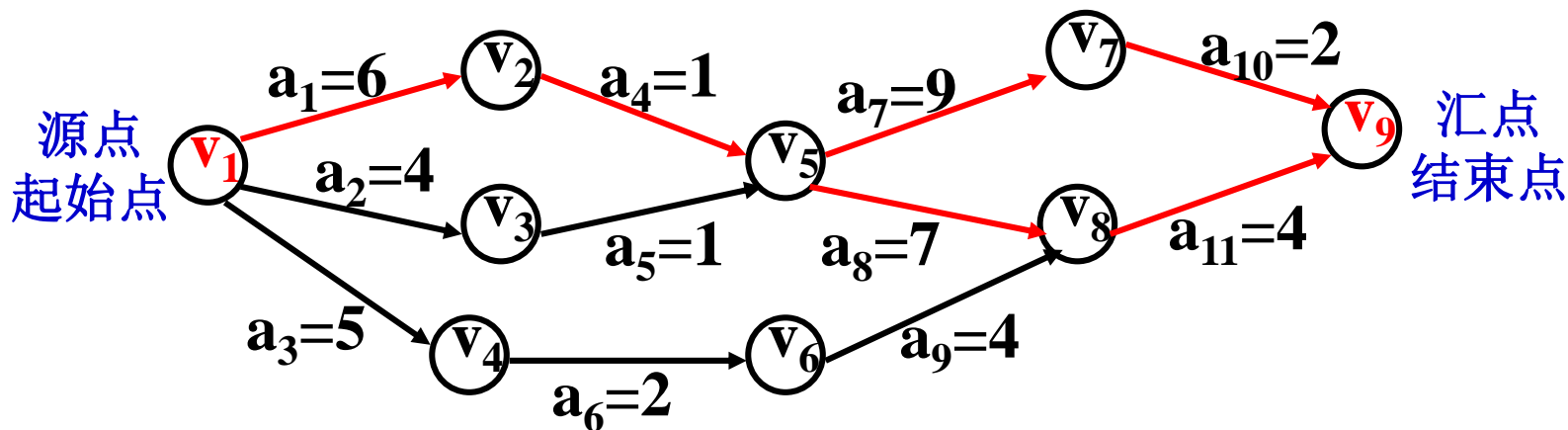




## 4.7 关键路径算法(cont.)

### ➤ 路径长度、关键路径、关键活动:

- **路径长度**: 是指从源点到汇点路径上所有活动的持续时间之和。
- **关键路径**: 在AOE网中, 由于有些活动可以并行, 所以**完成工程的最短时间是从源点到汇点的最大路径长度**。因此, 把从源点到汇点具有最大长度的路径称为**关键路径**。
- 一个AOE中, 关键路径可能不只一条。
- **关键活动**: 关键路径上的活动称为**关键活动**。





## 4.7 关键路径算法(cont.)

➔ 关键路径和关键活动性质分析——与计算关键活动有关的量

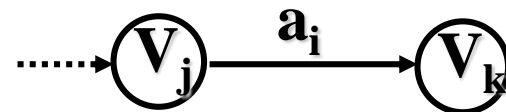
■ ①事件 $V_j$  的最早可能发生时间 $VE(j)$

● 是从源点 $V_1$  到顶点 $V_j$  的最长路径长度。

■ ②活动 $a_i$  的最早可能开始时间  $E(i)$

● 设活动 $a_i$  在边 $\langle V_j, V_k \rangle$ 上, 则 $E(i)$ 也是从源点 $V_1$ 到顶点 $V_j$  的最长路径长度。这是因为事件 $V_j$ 发生表明以 $V_j$ 为起点的所有活动 $a_i$ 可以立即开始。因此,

$$E(i) = VE(j) \dots\dots\dots (1)$$



有多个前驱顶点:

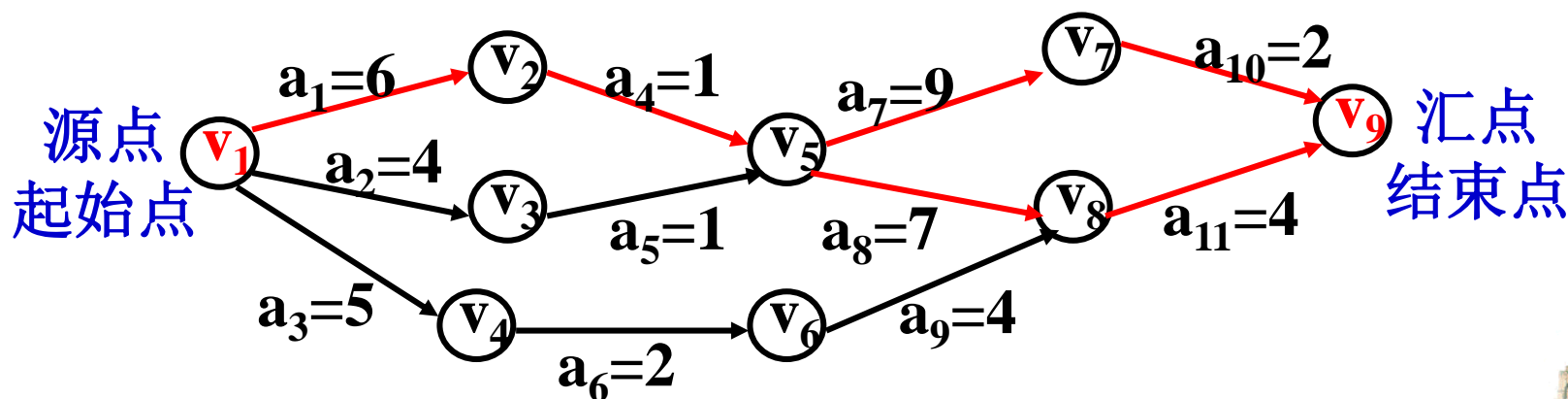
$$\begin{aligned}
 VE(v_5) &= \max \{ve(\text{前驱顶点}) + \text{前驱活动时间}\} \\
 &= \max \{6+1, 4+1\} = 7
 \end{aligned}$$

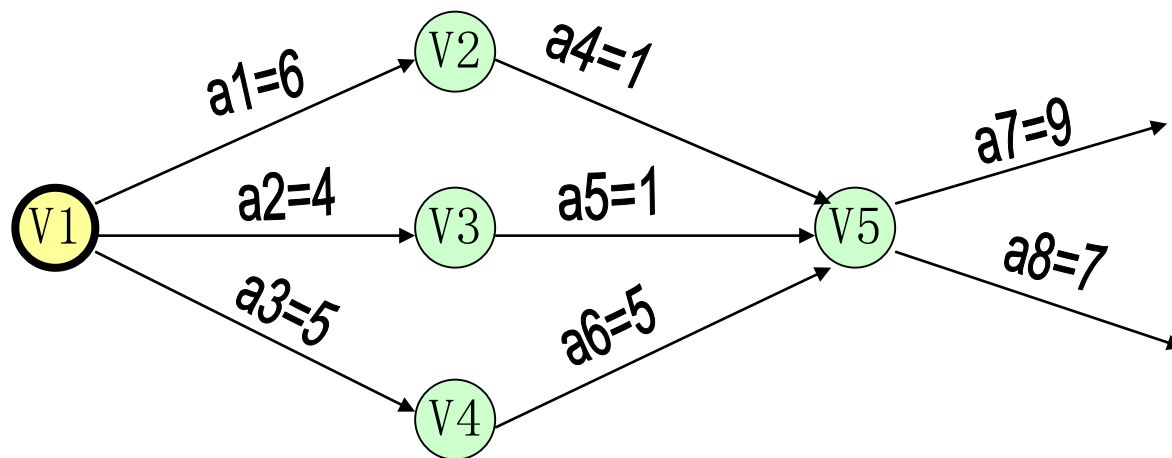
事件 $V_5$ 的最早发生时间

$$a_1 + a_4 = 6 + 1 = 7$$

活动 $a_7, a_8$ 的最早发生时间

$$E(7) = E(8) = 6 + 1 = 7$$





各顶点事件最早开始时间:

$$VE(v1)=0 \quad VE(v2)=6$$

$$VE(v3)=4 \quad VE(v4)=5$$

$$VE(v5)=10$$

各活动最早开始时间:

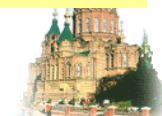
$$E(a1)=E(a2)=E(a3)=VE(v1)=0$$

$$E(a4)=VE(v2)=6$$

$$E(a5)=VE(v3)=4$$

$$E(a6)=VE(v4)=5$$

$$E(a7)=E(a8)=VE(v5)=10$$



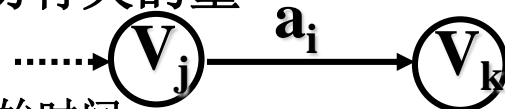


## 4.7 关键路径算法(cont.)

### ➤ 关键路径和关键活动性质分析-----与计算关键活动有关的量

#### ■ ③事件 $V_k$ 的最迟发生时间 $VL(k)$

- 在 $VE(n)$ 时刻完成的前提下，事件 $V_k$ 的允许的最迟开始时间。
- 一个事件最迟发生时间 $VL(k)$ 应该等于汇点的最早发生时间 $VE(n)$ 减去从 $V_k$ 到 $V_n$ 的最大路径长度。



#### ■ ④活动 $a_i$ 的最迟允许开始时间 $L(i)$

- 不影响工期的前提下，活动 $a_i$ 允许的最迟开始时间。
- 因为事件 $V_k$ 发生表明以 $V_k$ 为终点的入边所表示的所有活动均已完成，所以事件 $V_k$ 的最迟发生时间 $VL(k)$ 也是所有以 $V_k$ 为终点的入边 $\langle V_j, V_k \rangle$ 所表示的活动 $a_i$ 可以最迟完成时间。
- 显然，为不推迟工期，活动 $a_i$ 的最迟开始时间 $L(i)$ 应该是 $V_k$ 的最迟完成时间 $VL(k)$ 减去 $a_i$ 的持续时间，即

$$\blacklozenge L(i) = VL(k) - ACT[j][k] \dots\dots\dots(2)$$

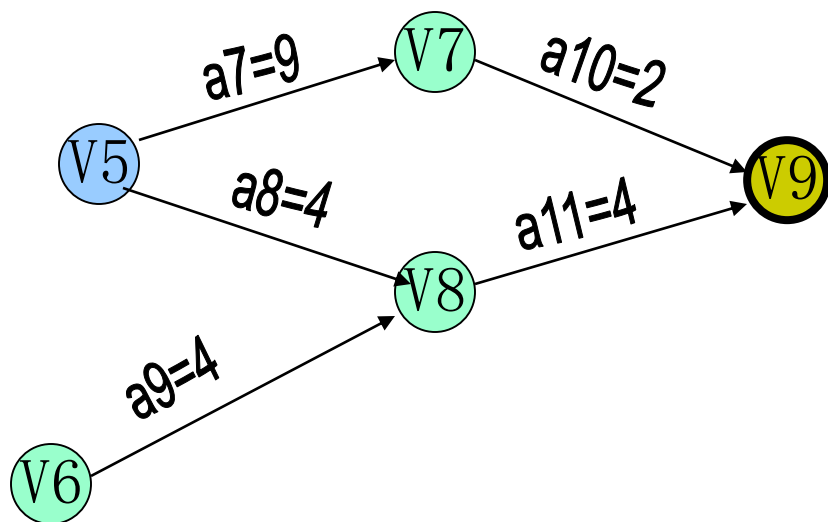
◆ 其中,  $ACT[j][k]$ 是活动 $a_i$ 的持续时间 ( $\langle V_j, V_k \rangle$ 上的权)。





不推迟整个工程完成的前提下，（顶点）事件 $V_i$ 允许的最迟开始时间 $VL(i)$ ：完成点（汇点） $V_n$ 的的最早发生时间 $VE(n)$ 减去 $V_k$ 到 $V_n$ 的最长路径长度。

（ $V_n$ 的的最早发生时间 $VE(n)$ 等于最迟开始时间 $VL(n)$ ）。



仅有一个后继顶点：

假定工程18天完成 ( $VE(V9)=18$ )，则：

$$VL(V9)=18$$

$$VL(V7)=VL(V9)-2=16$$

$$VL(v8)=VL(V9)-4=14$$

$$VL(v6)=VL(v8)-4=10$$

有多个后继顶点：

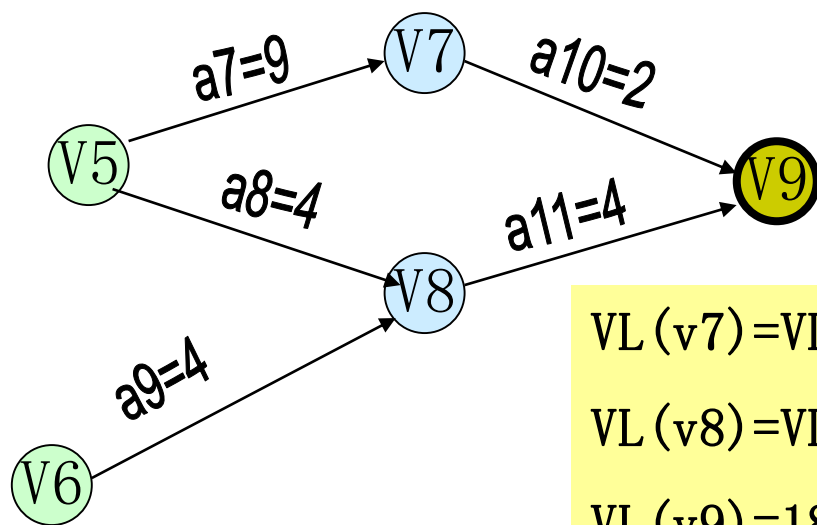
$$VL(v5)=\min\{VL(v7)-9, VL(v8)-4\}=\min\{7, 10\}=7$$





确定了顶点 $v_i$ 的最迟开始时间后，确定所有以 $v_i$ 为弧头的活动 $a_k$ 的最迟开始时间 $L(k)$ ：表示在不推迟整个工程完成的前提下，活动 $a_k$ 最迟必须开始的时间。

$$L(a_k) = VL(a_k \text{ 弧头对应顶点}) - \text{活动 } a_k \text{ 的持续时间}$$



$$VL(v_7) = VL(v_9) - 2 = 16$$

$$VL(v_8) = VL(v_9) - 4 = 14$$

$$VL(v_9) = 18$$

$$L(a_{11}) = VL(v_9) - 4 = 18 - 4 = 14$$

$$L(a_{10}) = VL(v_9) - 2 = 18 - 2 = 16$$

$$L(a_9) = VL(v_8) - 4 = 14 - 4 = 10$$

$$L(a_8) = VL(v_8) - 4 = 14 - 4 = 10$$

$$L(a_7) = VL(v_7) - 9 = 16 - 9 = 7$$

$L(i) - E(i)$  意味着完成活动 $a_i$ 的时间余量。

关键活动：  $L(i) = E(i)$  的活动。





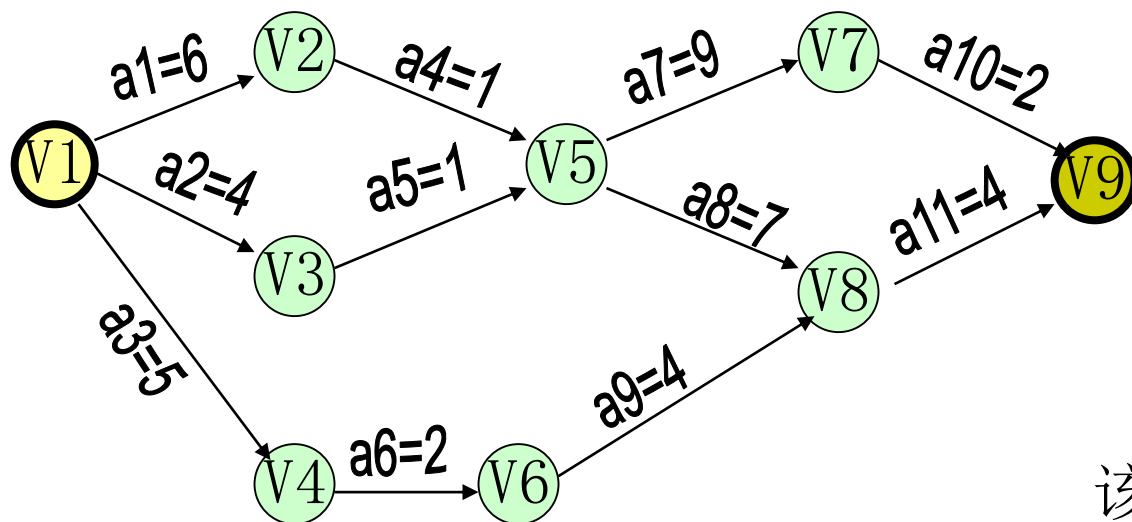
### 关键路径算法步骤:

(1) 从开始点V1出发, 令 $VE(1)=0$ , 按拓扑排序序列求其它各顶点的最早发生时间

$$VE(k) = \max \{VE(j) + act(<j, k>)\}$$

( $v_j$ 为以顶点 $v_k$ 为弧头的所有弧的弧尾

对应的顶点集合)



顶点	$VE(i)$	$VL(i)$
$v_1$	0	
$v_2$	6	
$v_3$	4	
$v_4$	5	
$v_5$	7, 5	
$v_6$	7	
$v_7$	16	
$v_8$	14, 11	
$v_9$	18, 18	

该表次序为一拓扑排序序列



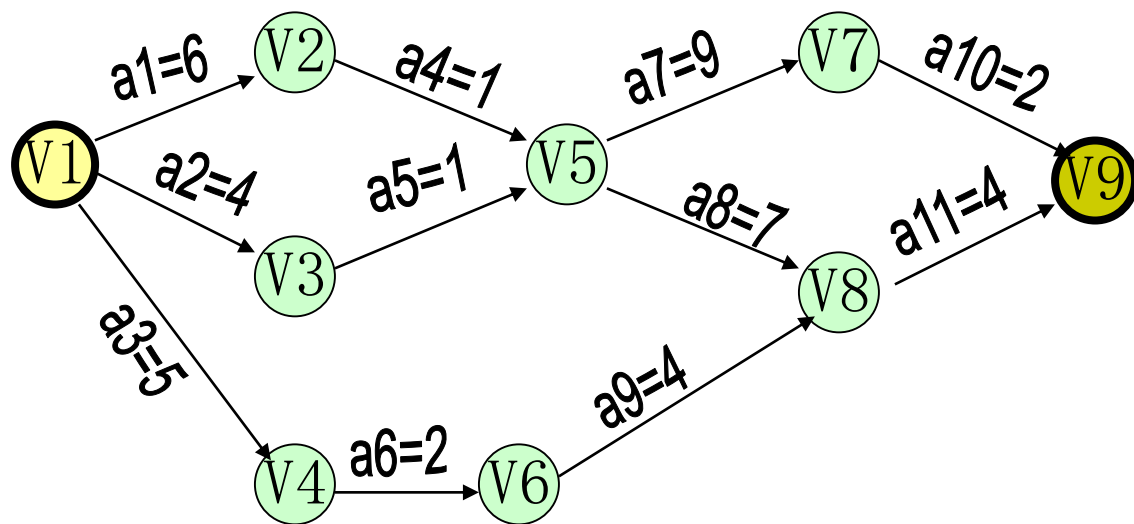


关键路径算法步骤:

(2) 从完成点 $v_n$ 出发, 令 $VL(n)=VE(n)$ , 按逆拓扑排序序列求其它各顶点的最迟发生时间

$$VL(j) = \min \{ VL(k) - ACT(<j, k>) \}$$

( $v_k$ 为以顶点 $v_j$ 为弧尾的所有弧的弧头对应的顶点集合)



顶点	VE(i)	VL(i)
$v_1$	0	0, 2, 3
$v_2$	6	6
$v_3$	4	6
$v_4$	5	8
$v_5$	7	7, 7
$v_6$	7	10
$v_7$	16	16
$v_8$	14	14
$v_9$	18	18



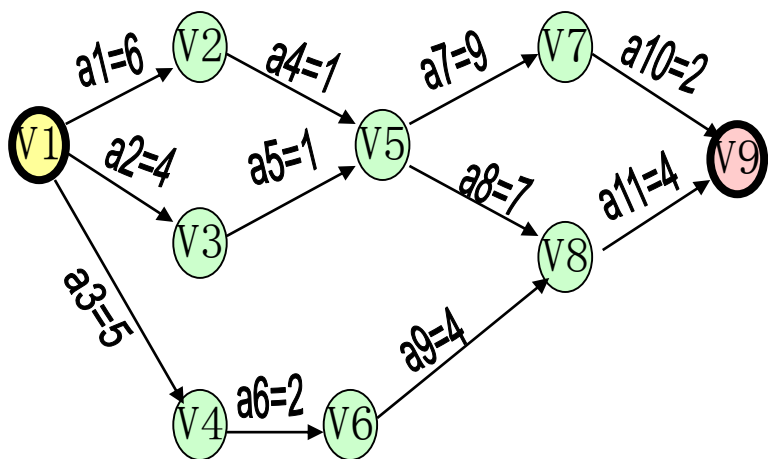




## 关键路径算法步骤:

(3) 求每一项活动 $a_i(v_j, v_k)$ :

$$E(i) = VE(j) \quad L(i) = VL(k) - ACT(a_i)$$



顶点	VE(i)	VL(i)
v <sub>1</sub>	0	0
v <sub>2</sub>	6	6
v <sub>3</sub>	4	6
v <sub>4</sub>	5	8
v <sub>5</sub>	7	7
v <sub>6</sub>	7	10
v <sub>7</sub>	16	16
v <sub>8</sub>	14	14
v <sub>9</sub>	18	18

活动	E(i)	L(i)	L(i)-E(i)
a <sub>1</sub>	0	0	0
a <sub>2</sub>	0	2	2
a <sub>3</sub>	0	3	3
a <sub>4</sub>	6	6	0
a <sub>5</sub>	4	6	2
a <sub>6</sub>	5	8	3
a <sub>7</sub>	7	7	0
a <sub>8</sub>	7	7	0
a <sub>9</sub>	7	10	3
a <sub>10</sub>	16	16	0
a <sub>11</sub>	14	14	0



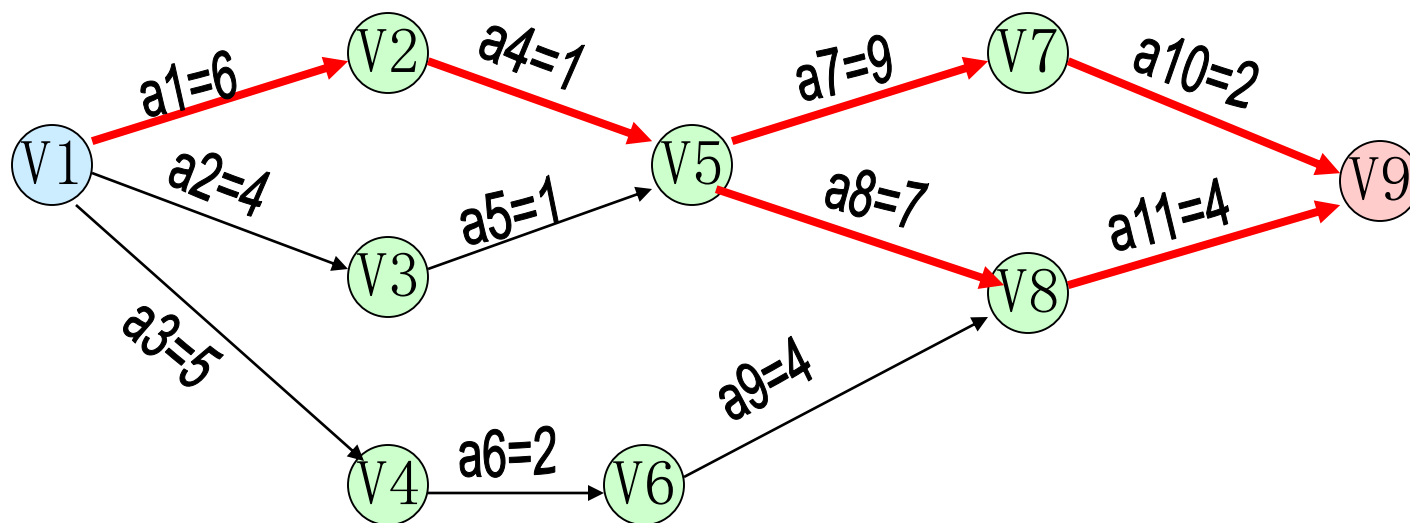


关键活动：选取 $E(i)=L(i)$ 的活动。

关键路径：

(1)  $v1 \rightarrow v2 \rightarrow v5 \rightarrow v7 \rightarrow v9$

(2)  $v1 \rightarrow v2 \rightarrow v5 \rightarrow v8 \rightarrow v9$

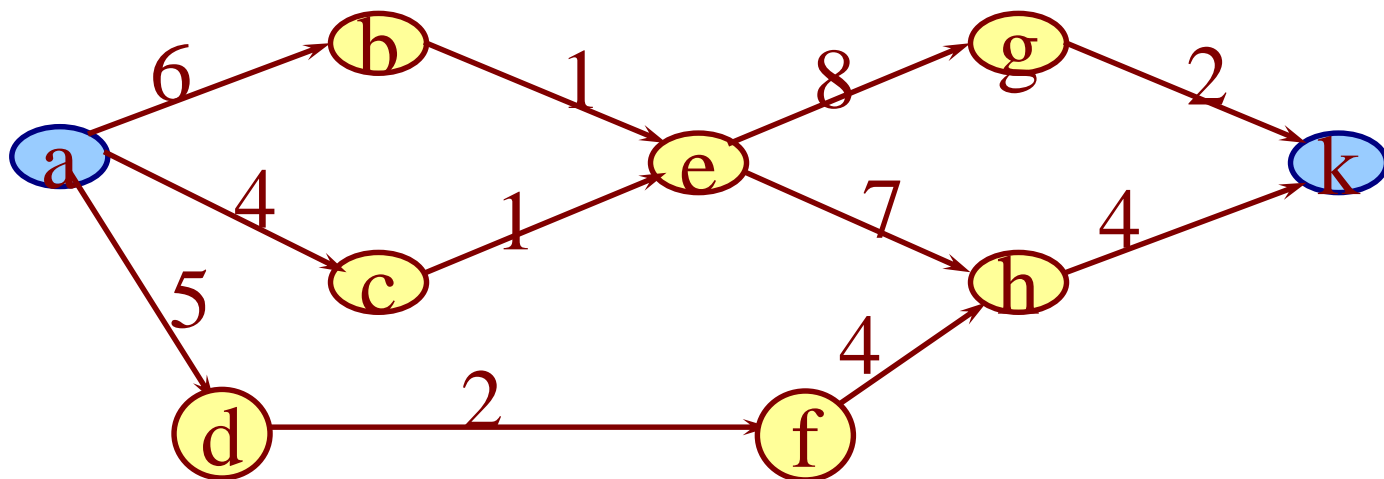




## 4.7 关键路径算法(cont.)

- AOE网可以用邻接矩阵或邻接表表示；矩阵中行下标*i*表示起点，列下标*j*表示终点， $ACT[i][j]>0$ 表示时间， $ACT[i][j]=-1$ 表示无边。

关键路径和关键活动分析计算示例：



	a	b	c	d	e	f	g	h	k
ve	0	6	4	5	7	7	15	14	18
vl	0	6	6	8	7	10	16	14	18





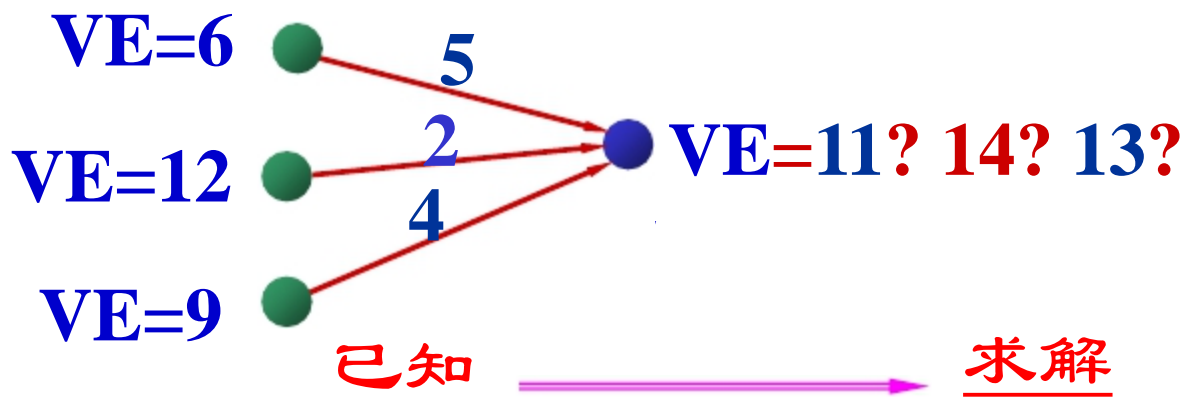
## 4.7 关键路径算法(cont.)

➡ 利用拓扑排序算法求关键路径和关键活动

- (1) 前进阶段：计算 $VE[j]$ 。从源点 $V_1$ 出发，令 $VE(1) = 0$ ，按拓扑序列次序求出其余各顶点事件的最早发生时间：

$$VE(j) = \max_{j \in T} \{ VE(i) + ACT[i][j] \}$$

- 其中 $T$ 是以顶点 $V_j$ 为尾的所有边的头顶点的集合( $2 \leq j \leq n$ )
- 如果网中有回路，不能求出关键路径则算法中止；否则转 (2)





## 4.7 关键路径算法(cont.)

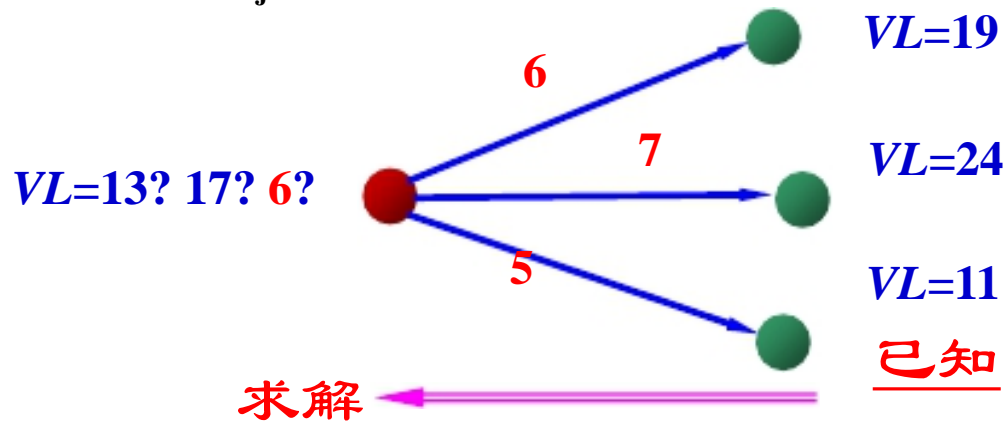
➔ 利用拓扑排序算法求关键路径和关键活动

- (2) 回退阶段: 计算 $VL[j]$ 从汇点 $V_n$ 出发, 令 $VL(n) = VE(n)$ , 按逆拓扑有序求其余各顶点的最晚发生时间(用逆邻接矩阵 $ACT$ 转置即可)

:

$$VL(j) = \min_{k \in S} \{ VL(i) - ACT[j][i] \}$$

- 其中 $S$ 是以顶点 $V_j$ 为头的所有边的尾顶点的集合( $2 \leq j \leq n-1$ )





### ■ (3) 计算 $E(i)$ 和 $L(i)$

求每一项活动 $a_i$ 的最早开始时间:

$$E(i) = VE(j)$$

求每一项活动 $a_i$ 的最晚开始时间:

$$L(i) = VL(k) - ACT[j][k]$$

### ■ (4) 若某条边满足 $E(i) = L(i)$ , 则它是关键活动。

◆为了简化算法, 可以在求关键路径之前已经对各顶点实现拓扑排序, 并按拓扑有序的顺序对各顶点重新进行了编号。

◆不是任意一个关键活动的加速一定能使整个工程提前。

◆想使整个工程提前, 要考虑各个关键路径上所有关键活动。





关键路径是利用进度模型时使用的一种进度网络分析技术。

关键路径沿着项目进度网络线进行正向与反向分析，从而计算出所有计划活动理论上的最早开始与完成日期、最迟开始与完成日期，不考虑任何资源限制。

由此计算而得到的最早开始与完成日期、最迟开始与完成日期不一定是项目的进度表，他们只不过指明计划活动在给定的活动持续时间、逻辑关系、时间提前量与滞后量，以及其他已知制约条件下应当安排的时间段与长短。

由于构成进度灵活余地的总时差可能为正、负或零值，最早开始与完成日期、最迟开始与完成日期的计算值可能在所有的路线上都相同，也可能不同。

在任何网络路线上，进度灵活余地的大小由最早与最迟日期两者之间正的差值决定，该差值叫做“总时差”。





### 作业：工程安排

一项工程由多道工序组成，按照施工过程的要求，这些工序之间，客观上有一个必须遵守的先后关系。对那些紧接在已知工序前的工序叫紧前工序，把在已知工序后边紧接的工序叫后项工序，只有已知工序的所有紧前工序都完成，已知工序才能开始施工。例如某工程的工序表如下：

序代号	紧前工序	完成时间	序代号	紧前工序	完成时间
A	-	6	F	C	2
B	-	2	G	D	3
C	A	3	H	B,E	4
D	A	5	I	H	2
E	A	3	J	F,G,I	2

一天中可以同时进行若干道工序。

编程实现：工程最少在几天内完成，并找出一种工程施工安排方案。







## 4.8 最短路径算法

### 最短路径(Shortest Path)问题

- 如果图中从一个顶点可以到达另一个顶点，则称这两个顶点间存在一条**路径**。
- 从一个顶点到另一个顶点间可能存在多条路径，而每条路径上经过的**边数**并不一定相同。
- 如果图是一个带权图，则路径长度为路径上各边的权值的总和，两个顶点间路径长度最短的那条路径称为两个顶点间的**最短路径**，其路径长度称为**最短路径长度**。
- 如何找到一条路径使得沿此路径上各边上的权值总和达到最小？
- 集成电路设计、GPS导航、路由选择、铺设管线等





## 4.8 最短路径算法(cont.)

### 问题解法

- 边上权值非负情形的单源最短路径问题

- — Dijkstra算法

- 所有顶点之间的最短路径问题

- — Floyd算法

### 边上权值非负情形的单源最短路径问题：

- **问题描述：** 给定一个带权有向图 $G=(V, E)$ 与源点  $v \in V$ ，求从  $v$  到 $G$ 中其它顶点的最短路径。限定各边上的权值大于或等于0。





## 4.8最短路径算法(cont.)

- 艾兹格·W·迪科斯彻 (Edsger Wybe Dijkstra, 1930年5月11日~2002年8月6日)荷兰人。计算机科学家，毕业就职于荷兰Leiden大学，早年钻研物理及数学，而后转为计算学。曾在1972年获得过素有计算机科学界的诺贝尔奖之称的图灵奖，之后，他还获得过1974年 **AFIPS Harry Goode Memorial Award**、1989年**ACM SIGCSE**计算机科学教育教学杰出贡献奖、以及2002年**ACM PODC**最具影响力论文奖。

- 1 提出“goto有害论”；结构程序设计之父
- 2 提出信号量和PV原语；
- 3 解决了有趣的“哲学家聚餐”问题；
- 4 最短路径优先算法（SPF）的创造者；
- 5 第一个Algol 60编译器的设计者和实现者；
- 6 THE操作系统的设计者和开发者；
- 7 提出银行家算法，解决操作系统中资源分配问题
- 8 与D. E. Knuth并称为我们这个时代最伟大的计算机科学家。

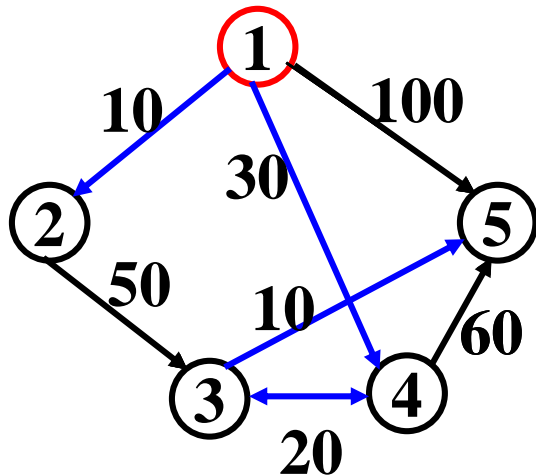




## 4.8 最短路径算法(cont.)

### ➔ Dijkstra算法的基本思想

- Dijkstra提出按路径长度的递增次序，逐步产生最短路径的贪心算法---Dijkstra算法（SPF算法）。
- 首先求出长度最短的一条最短路径,再参照它求出长度次短的一条最短路径，依次类推，直到从顶点 $v$  到其它各顶点的最短路径全部求出为止。



源点S	中间结点	终点	路径长度
1		2	1 0
1		4	3 0
1	4	3	5 0
1	4 3	5	6 0

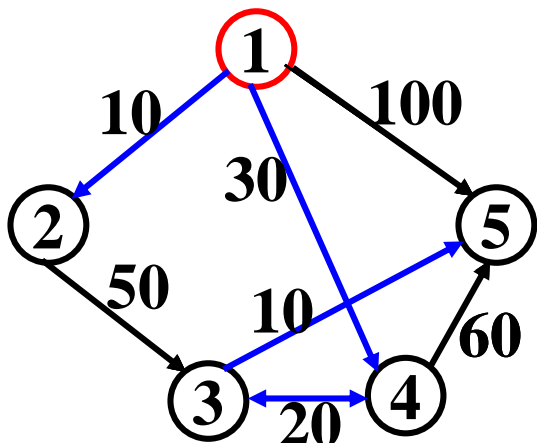




## 4.8 最短路径算法(cont.)

### ➡ Dijkstra算法的数据结构

- 假设带权有向图 $G=(V, E)$ ，其中 $V=\{1, 2, \dots, n\}$ ，顶点1为源点。图 $G$ 的存储结构：采用带权的邻接矩阵 $C$ 表示。
- 一维数组 $D[n]$ ：  $D[i]$ 表示源点1到顶点 $i$ 的当前最短路径长度，初始时， $D[i]=C[1][i]$ ；
- 一维数组 $P[n]$ ：  $P[i]$ 表示源点1到顶点 $i$ 的当前最短路径上，最后经过的顶点，初始时， $P[i]=1$ （源点）；
- $S[n]$ ： 存放源点和已生成的终点，其初态为只有一个源点 $v$



$$C = \begin{bmatrix} \infty & 10 & \infty & 30 & 100 \\ \infty & \infty & 50 & \infty & \infty \\ \infty & \infty & \infty & 20 & 10 \\ \infty & \infty & 20 & \infty & 60 \\ \infty & \infty & \infty & \infty & \infty \end{bmatrix}$$

D		10	$\infty$	30	100
P		1	1	1	1
S	T	F	F	F	F



## 4.8 最短路径算法(cont.)

### ➤ Dijkstra算法实现步骤:

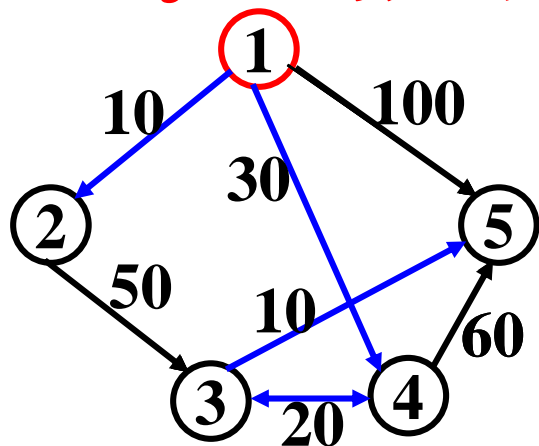
- 1. 将  $V$  分为两个集合  $S$ （最短路径已经确定的顶点集合）和  $V-S$ （最短路径尚未确定的顶点集合。初始时， $S=\{1\}$ ， $D[i]=C[1][i]$  ( $i=2,3,\dots,n$ )， $P[i]=1$ (源点， $i \neq 1$ )。
- 2. 从  $S$  之外即  $V-S$  中选取一个顶点  $w$ ，使  $D[w]$  最小（即选这样的  $w$ ， $D[w]=\min\{D[i] \mid i \in V-S\}$ ），于是从源点到达  $w$  只通过  $S$  中的顶点，且是一条最短路径（选定路径），并把  $w$  加入集合  $S$ 。
- 3. 调整  $D$  中记录的从源点到  $V-S$  中每个顶点的最短距离，即从原来的  $D[v]$  和  $D[w]+C[w][v]$  中选择最小值作为  $D[v]$  的新值，且  $P[v]=w$ 。
- 4. 重复2和3，直到  $S$  中包含  $V$  的所有顶点为止。结果数组  $D$  就记录了从源到  $V$  中各顶点的最短距离（数组  $P$  记录最短路径）。



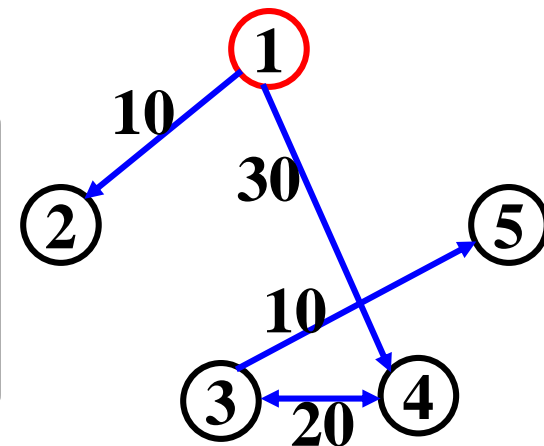


## 4.8 最短路径算法(cont.)

### Di jkstra算法示例



$$C = \begin{bmatrix} \infty & 10 & \infty & 30 & 100 \\ \infty & \infty & 50 & \infty & \infty \\ \infty & \infty & \infty & 20 & 10 \\ \infty & \infty & 20 & \infty & 60 \\ \infty & \infty & \infty & \infty & \infty \end{bmatrix}$$



循环	S	w	D[2]	D[3]	D[4]	D[5]	P[2]	P[3]	P[4]	P[5]
初态	{1}	-	10	$\infty$	30	100	1	1	1	1
1	{1,2}	2	10	60	30	100	1	2	1	1
2	{1,2,4}	4	10	50	30	90	1	4	1	4
3	{1,2,4,3}	3	10	50	30	60	1	4	1	3
4	{1,2,4,3,5}	5	10	50	30	60	1	4	1	3





## 4.8最短路径算法(cont.)

### ➡ Dijkstra算法的实现

```
void Dijkstra(GRAPH C, costtype D[n+1], int P[n+1], bool S[n+1])
```

```
{ for ( i=2 ; i<=n; i++ )
```

```
{   D[i]=C[1][i] ; S[i]=False ;P[i]=1;}
```

```
S [1]= True ;
```

```
for( i=1; i<=n-1; i++)
```

```
{ w=MinCost ( D, S ) ;
```

```
   S[w]=True ;
```

```
   for ( v=2 ; v<= n ; v++)
```

```
       if ( S[v]!=True )
```

```
       {   sum=D[w] + C[w][v] ;
```

```
           if (sum < D[v] ){D[v] = sum ; P[v]=w;}}
```

```
}
```

```
// 时间复杂度:  $O(n^2)$ 
```

```
costtype MinCost (D, S)
```

```
{
```

```
temp = INFINITY ;
```

```
w = 2 ;
```

```
for ( i=2 ; i<=n ; i++ )
```

```
    if (!S[i]&&D[i]<temp)
```

```
    { temp = D[i] ;
```

```
      w = i ;
```

```
    }
```

```
return w ;
```

```
}
```







## ➤ 普里姆 (Prim) 算法的实现

```
void Prim(Costtype C[n+1][n+1] )
{ costtype LOWCOST[n+1]; int CLOSEST[n+1]; int
  i,j,k; costtype min;
  for( i=2; i<=n; i++ )
  { LOWCOST[i] = C[1][i];   CLOSEST[i] = 1; }
  for( i = 2; i <= n; i++ )
  { min = LOWCOST[i];
    k = i;
    for( j = 2; j <= n; j++ )
      if ( LOWCOST[j] < min )
        { min = LOWCOST[j]; k=j; }
    cout << "(" << k << "," << CLOSEST[k] << ")"
    << endl;
    LOWCOST[k] = infinity;
    for ( j = 2; j <= n; j++ )
      if ( C[k][j] < LOWCOST[j] && LOWCOST[j] < infinity )
        { LOWCOST[j]=C[k][j]; CLOSEST[j]=k; }
  }
} /* 时间复杂度: O(|V|^2)
```

## ➤ Dijkstra算法的实现

```
void Dijkstra(GRAPH C, costtype D[n+1] , int P[n+1], bool
  S[n+1])
{ for ( i=2 ; i<=n; i++ )
  { D[i]=C[1][i] ; S[i]=False ;P[i]=1;}
  S [1]= True ;
  for( i=1; i<=n-1; i++ )
  { w=MinCost ( D, S ) ;
    S[w]=True ;
    for ( v=2 ; v<= n ; v++ )
      if ( S[v]!=True )
        { sum=D[w] + C[w][v] ;
          if (sum < D[v] ){D[v] = sum ; P[v]=w;}}
        }
  }
} // 时间复杂度: O (n^2)
```

```
costtype MinCost (D, S)
{
  temp = INFINITY ;
  w = 2 ;
  for ( i=2 ; i<=n ; i++ )
    if (!S[i]&&D[i]<temp)
      { temp = D[i] ;
        w = i ;
      }
  return w ;
}
```





### 贪心算法

贪心算法是从局部的最优出发。对于一个问题的最优解只能用穷举法得到时，用贪心法是寻找问题次优解的较好算法。

贪心法是一种改进了的分级处理方法。用贪心法设计算法的特点是一步一步地进行，根据某个优化测度（可能是目标函数，也可能不是目标函数），每一步都要保证能获得局部最优解。每一步只考虑一个数据，它的选取应满足局部优化条件。若下一个数据与部分最优解连在一起不再是可行解时，就不把该数据添加到部分解中，直到把所有数据枚举完，或者不能再添加为止。这种能够得到某种度量意义下的最优解的分级处理方法称为贪心法。

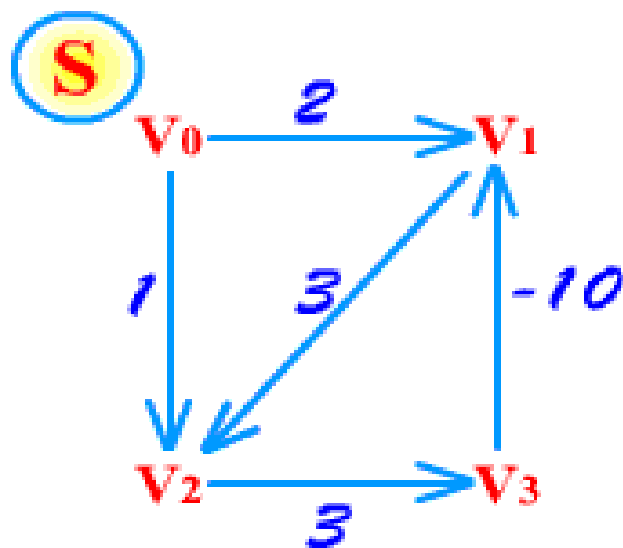
选择能产生问题最优解的最优度量标准是使用贪心法的核心问题。





## 负权最短路径问题

**Dijkstra**算法要求边的权值非负。事实上，一旦某些边的权值为负，那么**Dijkstra**算法就无法得出正确的最小路径长度。因为根据选取规则选取 $V$ 访问时，得到 $D_v$ ，但此时的路径可能没有经过顶点 $u$ ，而 $\text{weight}(\langle u, v \rangle) < 0$ ，如果 $D_u + \text{weight}(\langle u, v \rangle) < D_v$ ，那么 $D_v$ 就不是 $v$ 的最小路径长度。



如图， $D_0=0, D_0-2=1, D_0-1=2, D_0-2-3=4$ ，但从 $S$ 到 $V_1$ 还有其它路径：路径 $V_0-V_2-V_3-V_1$ 的长度为 $-6$ ；路径 $V_0-V_2-V_3-V_1-V_2-V_3-V_1$ 的长度为 $-10$ ，如此下去， $V_1$ 的最小路径长度是 $-\infty$ 。

出现这种情况的原因是图中出现了包括负权边的回路，我们称这种回路为负开销回路。





## 负权最短路径问题算法

在算法中，设置了一个队列，用来存放待访问顶点。并且增设数组 $scratch[]$ ，数组元素的初值均为0。当一个顶点 $V$ 入队时，它对应的 $scratch[v]$ 值加1；当该顶点出队时，其 $scratch[v]$ 值再加1。这就是说，若 $scratch[v]$ 为奇数，说明 $V$ 在队列中；若 $scratch[v]$ 为偶数，说明 $V$ 已出队。对于已经在队列中的顶点 $W$ ，当要改变 $D_w$ 时，无需令 $W$ 再出入队各一次，只须令

$$scratch[w]=scratch[w]+2$$

如此一来，逻辑上意味着 $W$ 出队一次，又入队一次。如果顶点 $W$ 存在不是负无穷的最短路径，那么它出入队次数均不超过 $n$ （ $n$ 是图中顶点个数）。因此，若 $scratch[w]>2n$ ，说明图中存在负开销回路，此时终止算法。





### 作业3:

相信大家都听说一个“百岛湖”的地方吧，百岛湖的居民生活在不同的小岛中，当他们想去其他的小岛时都要通过划小船来实现。现在政府决定大力发展百岛湖，发展首先要解决的问题当然是交通问题，政府决定实现百岛湖的全畅通！经过考察小组RPRush对百岛湖的情况充分了解后，决定在符合条件的小岛间建上桥，所谓符合条件，就是2个小岛之间的距离不能小于10米，也不能大于1000米。当然，为了节省资金，只要求实现任意2个小岛之间有路通即可。其中桥的价格为 100元/米。

已知小岛的个数为 $M \leq 100$ ， $(x,y)$ 代表每个小岛的坐标，这些坐标都是  $0 \leq x, y \leq 1000$  的整数。

求出建桥的最小花费，不能全场通，输出no.





## 4.8 最短路径算法(cont.)

### ➤ 其它最短路径问题及解法

#### ■ 单目标最短路径问题:

找出图中每个顶点 $v$  到某个指定结点 $c$  最短路径

- 只需每边取反?

#### ■ 单结点对间最短路径问题:

对于某对顶点 $u$  和  $v$ ,找出 $u$  到 $v$  的一条最短路径

- 以 $u$  为源点

#### ■ 所有顶点间的最短路径问题:

对图中每对顶点 $u$  和 $v$ , 找出 $u$  到 $v$  的最短路径

- 以每个顶点为源点
- 直接用Floyd算法

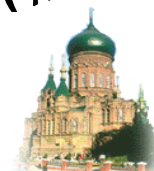




## 4.8 最短路径算法(cont.)

任意两个顶点之间的**最短路径**

- 问题描述：已知一个带权的有向图 $G=(V, E)$ ，对每一对顶点 $v_i, v_j \in V, (i \neq j)$ ，要求：求出 $v_i$ 与 $v_j$ 之间的**最短路径**和**最短路径长度**。
- 限制条件：不允许有负长度的环路。
- Floyd算法的基本想法：动态规划算法
  - 如果 $v_i$ 与 $v_j$ 之间有有向边，则 $v_i$ 与 $v_j$ 之间有一条路径，但不一定是最短的；也许经过某些中间点会使路径长度更短。
  - 经过哪些中间点会使路径长度**缩短**呢？经过哪些中间点会使路径长度**最短**呢？
    - 只需尝试在原路径中间**加入其它顶点**作为中间顶点。
  - 如何尝试？
    - 系统地在原路径中间加入每个顶点，然后不断地调整当前路径(和路径长度)即可。





## 4.8 最短路径算法(cont.)

■ 示例:

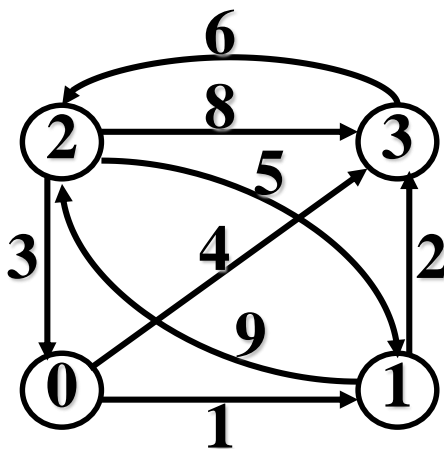
●  $\langle 2, 1 \rangle = 5$   $\langle 2, 0 \rangle + \langle 0, 1 \rangle = 4$   $a[2][1] = a[2][0] + a[0][1]$  调整

● 注意: 考虑  $v_0$  做中间点可能还会改变其它顶点间的距离:

$\langle 2, 0, 3 \rangle = 7$   $\langle 2, 3 \rangle = 8$   $a[2][3] = a[2][0] + a[0][3]$

●  $\langle 2, 3 \rangle$ :  $\langle 2, 0 \rangle + \langle 0, 3 \rangle$ :  $\langle 2, 0 \rangle + \langle 0, 1 \rangle + \langle 1, 3 \rangle = \langle 2, 0, 1, 3 \rangle$   $a[2][3] = 6$  调整

● 注意: 有时加入中间顶点后的路径比原路径长 保持



$$A = \begin{matrix} & \begin{matrix} 0 & 1 & 2 & 3 \end{matrix} \\ \begin{matrix} 0 \\ 1 \\ 2 \\ 3 \end{matrix} & \begin{pmatrix} \infty & 1 & \infty & 4 \\ \infty & \infty & 9 & 2 \\ 3 & 5 & \infty & 8 \\ \infty & \infty & 6 & \infty \end{pmatrix} \end{matrix} \begin{matrix} 0 \\ 1 \\ 2 \\ 3 \end{matrix}$$







## 4.8 最短路径算法(cont.)

### ➤ Floyd算法的基本思想:

- 假设求顶点  $v_i$  到顶点  $v_j$  的最短路径。如果从  $v_i$  到  $v_j$  存在一条长度为  $C[i][j]$  的路径, 该路径不一定是最短路径, 尚需进行  $n$  次试探。
- 首先考虑路径  $(v_i, v_0, v_j)$  是否存在。如果存在, 则比较  $(v_i, v_j)$  和  $(v_i, v_0, v_j)$  的路径长度取长度较短者为从  $v_i$  到  $v_j$  的中间顶点的序号不大于0的最短路径。
- 假设在路径上再增加一个顶点  $v_1$ , 也就是说, 如果  $(v_i, \dots, v_1)$  和  $(v_1, \dots, v_j)$  分别是当前找到的中间顶点的序号不大于0的最短路径, 那么  $(v_i, \dots, v_1, \dots, v_j)$  就是有可能是从  $v_i$  到  $v_j$  的中间顶点的序号不大于1的最短路径。将它与已经得到的从  $v_i$  到  $v_j$  中间顶点序号不大于0的最短路径相比较, 从中选出中间顶点的序号不大于1的最短路径, 再增加一个顶点  $v_2$ , 继续进行试探。
- 一般情况下, 若  $(v_i, \dots, v_k)$  和  $(v_k, \dots, v_j)$  分别是从小  $v_i$  到  $v_k$  和从  $v_k$  到  $v_j$  的中间顶点序号不大于  $k-1$  的最短路径, 则将  $(v_i, \dots, v_k, \dots, v_j)$  和已经得到的从  $v_i$  到  $v_j$  且中间顶点序号不大于  $k-1$  的最短路径相比较, 其长度较短者便是从  $v_i$  到  $v_j$  的中间顶点的序号不大于  $k$  的最短路径。





- 假设已求出 $A_{(k-1)}(i, j)$  ( $1 \leq i, j \leq n$ ), 怎样求出  $A_{(k)}(i, j)$ ?
- 如果从顶点 $i$ 到顶点 $j$ 的最短路径不经过顶点 $k$ , 则由 $A_{(k)}(i, j)$ 的定义可知, 从 $i$ 到 $j$ 的中间顶点序号不大于 $k$ 的最短路径长度就是 $A_{(k-1)}(i, j)$ , 即 $A_{(k)}(i, j) = A_{(k-1)}(i, j)$ 。
- 如果从顶点 $i$ 到顶点 $j$ 的最短路径经过顶点 $k$ , 则这样的一条路径是由 $i$ 到 $k$ 和由 $k$ 到 $j$ 的两条路径所组成。由于 $A_{(k-1)}(i, k) + A_{(k-1)}(k, j) < A_{(k-1)}(i, j)$ , 则 $A_{(k)}(i, j) = A_{(k-1)}(i, k) + A_{(k-1)}(k, j)$ 。





## 4.8 最短路径算法(cont.)

### ➡ Floyd算法的数据结构

#### ■ 图的存储结构:

- 带权的有向图采用邻接矩阵 $C[n][n]$ 存储

#### ■ 数组 $A[n][n]$ :

- 存放在迭代过程中求得的最短路径长度。迭代公式:

$$\begin{cases} A_0[i][j] = C[i][j] \\ A_k[i][j] = \min\{ A_{k-1}[i][j], A_{k-1}[i][k] + A_{k-1}[k][j] \} \end{cases} \quad 0 \leq k \leq n-1$$

#### ■ 数组 $P[n][n]$ :

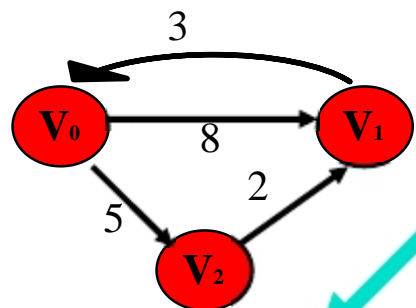
- 存放从  $v_i$  到  $v_j$  求得的最短路径。初始时,  $P[i][j] = -1$



## Floyd 算法的求解过程



实例及求解过程:



设  $C$  为  $n$  行  $n$  列的代价矩阵,  $c[i, j]$  为  $i \rightarrow j$  的权值。如果  $i=j$ ; 那么  $c[i, j]=0$ 。如果  $i$  和  $j$  之间无有向边;

则  $c[i, j]=\infty$

$$C = \begin{matrix} & \begin{matrix} 0 & 1 & 2 \end{matrix} \\ \begin{matrix} 0 \\ 1 \\ 2 \end{matrix} & \begin{bmatrix} 0 & 8 & 5 \\ 3 & 0 & \infty \\ \infty & 2 & 0 \end{bmatrix} \end{matrix}$$

1、使用  $n$  行  $n$  列的矩阵  $A$  用于计算最短路径。

初始时,  $A[i, j] = c[i, j]$

2、进行  $n$  次迭代

在进行第  $k$  次迭代时, 将使用如下的公式:

$$A_k[i, j] = \min \begin{cases} A_{k-1}[i, j] \\ A_{k-1}[i, k] + A_{k-1}[k, j] \end{cases}$$

注意: 第  $k$  次迭代时, 针对结点  $k$  进行。原  $A_{k-1}$  矩阵的第  $k$  行, 第  $k$  列保持不变。左上至右下的对角线元素也不变。

$$\begin{bmatrix} 0 & 8 & 5 \\ 3 & 0 & \infty \\ \infty & 2 & 0 \end{bmatrix} \rightarrow \begin{bmatrix} 0 & 8 & 5 \\ 3 & 0 & 8 \\ \infty & 2 & 0 \end{bmatrix}$$

$A$  初值

$A_0$

$$\begin{bmatrix} 0 & 8 & 5 \\ 3 & 0 & 8 \\ 5 & 2 & 0 \end{bmatrix} \rightarrow \begin{bmatrix} 0 & 7 & 5 \\ 3 & 0 & 8 \\ 5 & 2 & 0 \end{bmatrix}$$

$A_1$

实例中: 注意:  $k = 0 \sim (n-1)$ 。如在由  $A$  初值得到  $A_0$  时

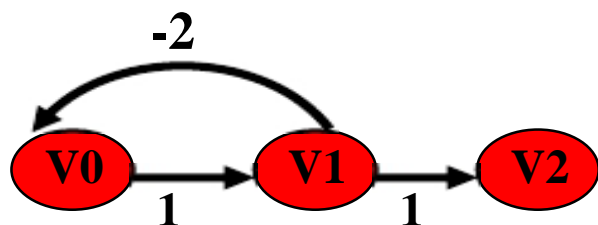
原  $A$  初值的第  $0$  行, 第  $0$  列不变, 仍体现在  $A_0$  中。





## Floyd 算法的求解过程

- 应用范围:无负长度的圈, 可有负长度的边。



$$\begin{bmatrix} 0 & 1 & \infty \\ -2 & 0 & 1 \\ \infty & \infty & 0 \end{bmatrix}$$

A初值

$$\begin{bmatrix} 0 & 1 & 2 \\ -2 & 0 & 1 \\ \infty & \infty & 0 \end{bmatrix}$$

A1

$$\begin{bmatrix} 0 & 1 & \infty \\ -2 & 0 & 1 \\ \infty & \infty & 0 \end{bmatrix}$$

A0

$A1[0, 2] \neq 2$   
 $\text{MIN} \{ A0[0, 2],$   
 $\quad A0[0, 1]$   
 $\quad + A0[1, 2]$   
 $\quad \}$

因为  $V0 \rightarrow V1 \rightarrow V0$   
可以有許多圈。





## 4.8 最短路径算法(cont.)

### → Floyd算法的实现

```
void Floyd( costtype A[][], costtype C[][], int P[][], int n)
{ for ( i = 0; i < n; i++ )
    for ( j = 0; j < n; j++ )
        { A[i][j] = C[i][j];
          P[i][j] = -1; }
    for ( k = 0; k < n; k++ )
        for ( i = 0; i < n; i++ )
            for ( j = 0; j < n; j++ )
                if ( A[i][k] + A[k][j] < A[i][j] )
                    { A[i][j] = A[i][k] + A[k][j];
                      P[i][j] = k; }
}
```

/\* 时间复杂度:  $O(n^3)$  \*/

### Warshall算法

求有向图邻接矩阵C的传递闭包D

$A[i][j] = A[i][j] \cup (A[i][k] \cap A[k][j])$ ;

可以判定有向图任意两点间是否存在有向路。





## 4.8 最短路径算法(cont.)

➤ 用  $path$  数组记录 经过路径 其实  $path$  的定义如下:

$path[i][j] = k$  表示最短路径  $i \rightarrow \dots \rightarrow j$  和  $j$  的直接前驱为  $k$  即是:  $i \rightarrow \dots \rightarrow k \rightarrow j$

比如:

$1 \rightarrow 5 \rightarrow 4$ ,  $4 \rightarrow 3 \rightarrow 6$  此时  $path[1][6] = 0$ ;  $0$  表示  $1 \rightarrow 6$  不通, 当引入结点  $k = 4$  此时有  $1 \rightarrow 5 \rightarrow 4 \rightarrow 3 \rightarrow 6$  显然有  $path[1][6] = 3 = path[4][6] = path[k][6]$

➤ 于是有  $path[i][j] = path[k][j]$

➤ 对于  $1 \rightarrow 5$  相邻边 我们可以在初始化时候 有  $path[1][5] = 1$ ;

➤ 如是对于 最短路径  $1 \rightarrow 5 \rightarrow 4 \rightarrow 3 \rightarrow 6$  有  $path[1][6] = 3$ ;  $path[1][3] = 4$ ;  
 $path[1][4] = 5$ ;  $path[1][5] = 1$  如此逆推可以得到最短路径记录值





此算法由Robert W. Floyd（罗伯特·弗洛伊德）于1962年发表在“Communications of the ACM”上。同年Stephen Warshall（史蒂芬·沃舍尔）也独立发表了这个算法。Robert W. Floyd这个牛人是朵奇葩，他原本在芝加哥大学读的文学，但是因为当时美国经济不太景气，找工作比较困难，无奈之下到西屋电气公司当了一名计算机操作员，在IBM650机房值夜班，并由此开始了他的计算机生涯。

此外他还和J. W. J. Williams（威廉姆斯）于1964年共同发明了著名的堆排序算法HEAPSORT。







## 4.8 最短路径算法(cont.)

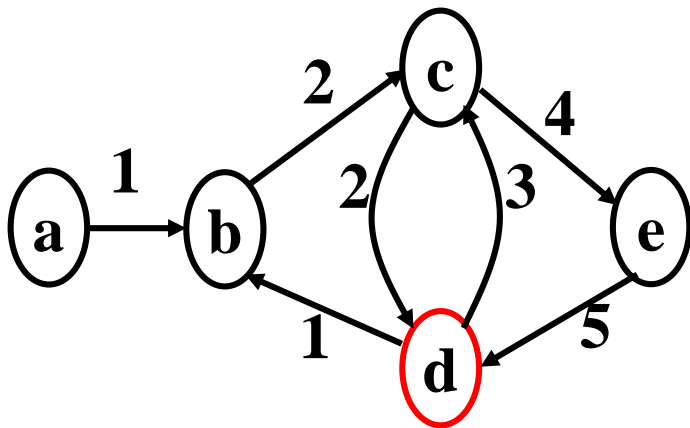
### Floyd算法的应用----求有向图的中心点

#### 顶点的偏心度:

- 设 $G=(V, E)$ 是一个带权有向图,  $D[i][j]$ 表示从 $i$ 到 $j$ 的最短距离。对任意一个顶点 $k$ ,  $E(k) = \max\{ d[i][k] \mid i \in V \}$ 称作顶点 $k$ 的**偏心度**。

#### 图 $G$ 的中心点:

- 称具有最小偏心度的顶点为图 $G$ 的**中心点**。



最短路径矩阵  $D$

	a	b	c	d	e
a	0	1	3	5	7
b	$\infty$	0	2	4	6
c	$\infty$	3	0	2	4
d	$\infty$	1	3	0	7
e	$\infty$	6	8	5	0

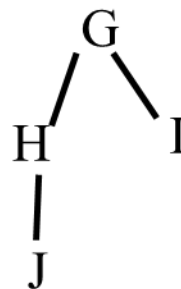
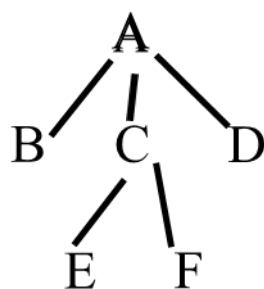
顶点	偏心度
a	$\infty$
b	6
c	8
d	5
e	7



思考题：森林采用带度数的后序遍历序列存储，设计算法将其转换成二叉树形式，二叉树的存储方案为 `[child][data][sibling]`。

森林的存储形式如下：

data	B	E	F	C	D	A	J	H	I	G
degree	0	0	0	2	0	3	0	1	0	2



算法思想：结点度数 $m$ 为0，则进栈； $m$ 不为0则从栈中弹出 $m$ 个结点作为该结点的子树，生成的新子树进栈；重复执行，直到数组中所有结点处理完毕，栈中的结点就是森林中所有树的根。





```
struct Node
{ char info ; int degree;}
Btree convert(Node *nodes , int size)
{ Btree cur,temp1,temp2;
  Stack s;
  for(int i=0;i<size;i++)
  {   if(nodes[i].degree ==0)
      Push(nodes[i].info, s);
      else
      {   cur=new Btree;
          cur->data=nodes[i].info;
          temp2=Pop(s);//temp2 指向栈顶元素
          for(int j=1;j<nodes[i].degree;j++)
              {temp1=Pop(s); temp1->sibling=temp2; temp2=temp1;}
          cur->child=temp2;
          Push(cur,s);
      }
  }
  temp2=Pop(s);//弹出森林中最后一个树的根
  while(!Empty(s))
  {   cur=Pop(s);
      cur->sibling=temp2;
      temp2=cur;
  }
  return temp2;
}
```





## 本章小结

