

# 算法设计与分析 Algorithms Design & Analysis

## 第十讲：图的算法

1

## 图(Graphs)

- Applications that involve not only a set of items, but also the connections between them (物体及物体之间的连接关系)



Maps(地图)



Schedules(日程表)



Computer networks  
(计算机网络)



Hypertext(超文本)



Circuits(电路板)

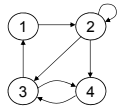
2

## 图的背景知识(Graphs – Background)

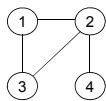
**Graphs = a set of nodes (vertices) with edges (links) between them.**(节点+边)

Notations:(表示方法)

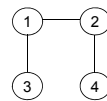
- $G = (V, E)$  – graph (图)
- $V$  = set of vertices  $|V| = n$  (节点和节点数目)
- $E$  = set of edges  $|E| = m$  (边和边的数目)



Directed  
Graph (有向图)



Undirected  
Graph (无向图)

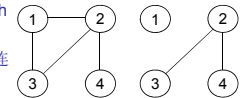


Acyclic  
Graph (无回路图)

3

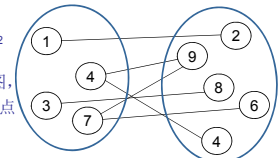
## 图的其它类型(Other Types of Graphs)

- A graph is **connected** if there is a path between every two vertices (如果任何两个顶点之间存在一个通路,则称图是连接的)



Connected Not connected

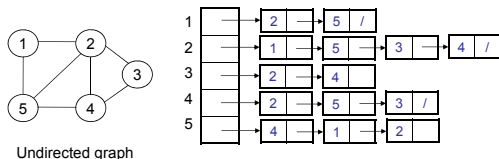
- A **bipartite graph** is an undirected graph  $G = (V, E)$  in which  $V = V_1 + V_2$  and there are edges only between vertices in  $V_1$  and  $V_2$  (二分图: 无向图, 由  $V_1$  和  $V_2$  组成, 只在  $V_1$  和  $V_2$  之间的顶点存在边)



4

## 图的表示方法 (Graph Representation)

- Adjacency list representation of  $G = (V, E)$** (邻接表)
  - An array of  $|V|$  lists, one for each vertex in  $V$ (顺序存储的顶点表)
  - Each list  $Adj[u]$  contains all the vertices  $v$  such that there is an edge between  $u$  and  $v$  (边表: 一个顶点的边表的每个表目对应与该顶点相关联的一条边)
    - $Adj[u]$  contains the vertices adjacent to  $u$  (in arbitrary order) (顺序无关)
  - Can be used for both directed and undirected graphs(适合有向或无向图)



Undirected graph

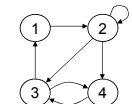
5

## 邻接表特性(Properties of Adjacency-List Representation)

- Sum of the lengths of all the adjacency lists (邻接表长度的和→边的数目)

- Directed graph: (有向图: 边出现一次)

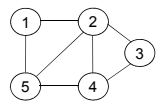
- Edge  $(u, v)$  appears only once in  $u$ 's list  $|E|$



Directed graph

- Undirected graph: (无向图: 边出现两次)

- $u$  and  $v$  appear in each other's adjacency lists:  
edge  $(u, v)$  appears twice  $2|E|$

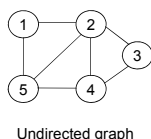


Undirected graph

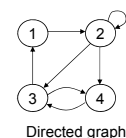
6

## 邻接表特性 (Properties of Adjacency-List Representation)

- **Memory required (空间需求)**
  - $\Theta(V + E)$
- **Preferred when (适合  $|E| \ll |V|^2$ )**
  - the graph is sparse:  $|E| \ll |V|^2$
- **Disadvantage(缺陷)**
  - no quick way to determine whether there is an edge between node  $u$  and  $v$  (不容易判断顶点 $u$ 和 $v$ 之间是否存在边)
- **Time to list all vertices adjacent to  $u$ : (罗列与邻接顶点的时间开销)**
  - $\Theta(\text{degree}(u))$
- **Time to determine if  $(u, v) \in E$ : (判断 $(u, v) \in E$ 的时间)**
  - $O(\text{degree}(u))$



Undirected graph

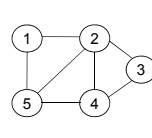


Directed graph

7

## 图的表示方法(Graph Representation)

- **Adjacency matrix representation of  $G = (V, E)$  (相邻矩阵)**
  - Assume vertices are numbered  $1, 2, \dots, |V|$  (顶点序号)
  - The representation consists of a matrix  $A_{|V| \times |V|}$ :
    - $a_{ij} = \begin{cases} 1 & \text{if } (i, j) \in E \text{ (元素值的确定)} \\ 0 & \text{otherwise} \end{cases}$



Undirected graph

	1	2	3	4	5
1	0	1	0	0	1
2	1	0	1	1	1
3	0	1	0	1	0
4	0	1	1	0	1
5	1	1	0	1	0

For undirected graphs matrix  $A$  is symmetric:

$$a_{ij} = a_{ji} \\ A = A^T \text{ (无向图的相邻矩阵的对称性)}$$

8

## 相邻矩阵的特性(Properties of Adjacency Matrix Representation)

- **Memory required (空间需求)**
  - $\Theta(V^2)$ , independent on the number of edges in  $G$  (与边无关)
- **Preferred when (适合情况)**
  - The graph is dense  $|E|$  is close to  $|V|^2$  ( $|E|$  接近  $|V|^2$ )
  - We need to quickly determine if there is an edge between two vertices (需要快速确定两个顶点之间是否存在边)
- **Time to list all vertices adjacent to  $u$ : (罗列与 $u$ 邻接顶点的时间)**
  - $\Theta(V)$
- **Time to determine if  $(u, v) \in E$ : (判断 $(u, v) \in E$ 的时间)**
  - $\Theta(1)$

9

## 有权图(Weighted Graphs)

- **Weighted graphs = graphs for which each edge has an associated weight  $w(u, v)$  (图的边被赋予了权值)**
  - $w: E \rightarrow \mathbb{R}$ , weight function (加权函数)
- **Storing the weights of a graph (权的存储)**
  - **Adjacency list: (邻接表)**
    - Store  $w(u, v)$  along with vertex  $v$  in  $u$ 's adjacency list (在边表条目中)
  - **Adjacency matrix: (相邻矩阵)**
    - Store  $w(u, v)$  at location  $(u, v)$  in the matrix (在元素中)

10

## 图的遍历(Searching in a Graph)

- **Graph searching = systematically follow the edges of the graph so as to visit the vertices of the graph (系统的沿着图的边访问图的顶点)**
- **Two basic graph searching algorithms: (两种基本搜索算法)**
  - Breadth-first search(广度优先搜索)
  - Depth-first search(深度优先搜索)
  - The difference between them is in the order in which they explore the unvisited edges of the graph (两种搜索方法的区别在于对图中未访问边的搜索顺序)
- **Graph algorithms are typically elaborations of the basic graph-searching algorithms (搜索是图的算法的基础, 其他有关图的算法基本上是在此基础上的精细化)**

11

## 广度优先搜索(Breadth-First Search (BFS))

- **Input: (输入)**
  - A graph  $G = (V, E)$  (directed or undirected) (图, 有向或无向)
  - A source vertex  $s \in V$  (源顶点)
- **Goal: (目的)**
  - Explore the edges of  $G$  to "discover" every vertex reachable from  $s$ , taking the ones closest to  $s$  first (从临近源顶点 $s$ 最近的顶点开始, 通过对图 $G$ 的边的探索发现从源顶点 $s$ 能够抵达的每个顶点)
- **Output: (输出)**
  - $d[v]$  = distance (smallest # of edges) from  $s$  to  $v$ , for all  $v \in V$  (从 $s$ 到 $v$ 的距离)
  - A "breadth-first tree" rooted at  $s$  that contains all reachable vertices (广度优先树: 以 $s$ 为根, 包含所有可以抵达的顶点)

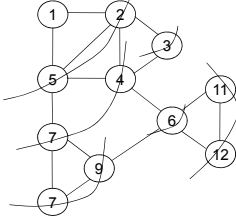
12

## 广度优先搜索 (Breadth-First Search)

- Discover vertices in increasing order of distance from the source  $s$  – search in breadth not depth

(广度优先: 距离源顶点的距离逐渐增加来搜索相连接的其他顶点)

- Find all vertices at 1 edge from  $s$ , then all vertices at 2 edges from  $s$ , and so on (距离递增)

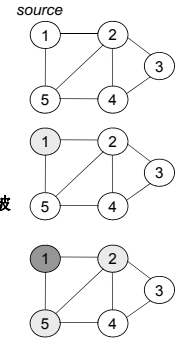


13

## 广度优先搜索(Breadth-First Search)

- Keeping track of progress: (过程跟踪)

- Color each vertex in either **white**, **gray** or **black** (用白、灰和黑来标记顶点)
- Initially, all vertices are **white** (最初为白)
- When being discovered a vertex becomes **gray** (刚搜索到, 灰)
- After discovering all its adjacent vertices the node becomes **black** (所有邻接顶点被搜索完, 黑)
- Use FIFO queue  $Q$  to maintain the set of gray vertices (先进先出队列来保存灰色顶点)

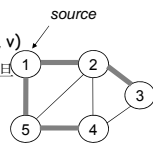


14

## 广度优先树(Breadth-First Tree)

- BFS constructs a breadth-first tree(BFS形成BFT)

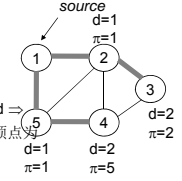
- Initially contains the root (source vertex  $s$ ) (源节点为树根)
- When vertex  $v$  is discovered while scanning the adjacency list of a vertex  $u \Rightarrow$  vertex  $v$  and edge  $(u, v)$  are added to the tree (在搜索顶点 $u$ 的邻接顶点时一旦发现了顶点 $v$ , 便将顶点 $v$ 和边 $(u, v)$ 加入到树中)
- $u$  is the **predecessor (parent)** of  $v$  in the breadth-first tree (在BFT中,  $u$ 是 $v$ 的父节点)
- A vertex is discovered only once  $\Rightarrow$  it has at most one parent (一个顶点只被发现一次, 因此只有一个父节点)



15

## BFS附加数据结构 (BFS Additional Data Structures)

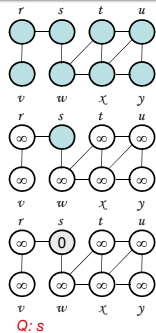
- $G = (V, E)$  represented using adjacency lists (邻接表法表示图)
- $color[u]$  – the color of the vertex for all  $u \in V$  (顶点颜色)
- $\pi[u]$  – predecessor of  $u$  (顶点 $u$ 的父顶点)
  - If  $u = s$  (root) or node  $u$  has not yet been discovered  $\Rightarrow \pi[u] = NIL$  (如果顶点是根顶点或没有被发现, 父顶点为空)
- $d[u]$  – the distance from the source  $s$  to vertex  $u$  (距离, 从 $s$ 到 $u$ )
- Use a FIFO queue  $Q$  to maintain the set of gray vertices (应用FIFO队列 $Q$ 存储灰色顶点)



16

## BFS算法: BFS( $V, E, s$ )

- for each  $u \in V - \{s\}$
- do  $color[u] \leftarrow WHITE$
- $d[u] \leftarrow \infty$
- $\pi[u] = NIL$
- $color[s] \leftarrow GRAY$
- $d[s] \leftarrow 0$
- $\pi[s] = NIL$
- $Q \leftarrow \emptyset$
- $Q \leftarrow ENQUEUE(Q, s)$

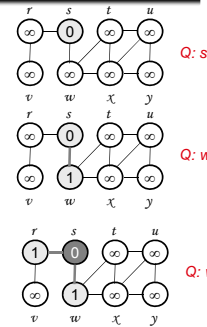


$Q: s$

17

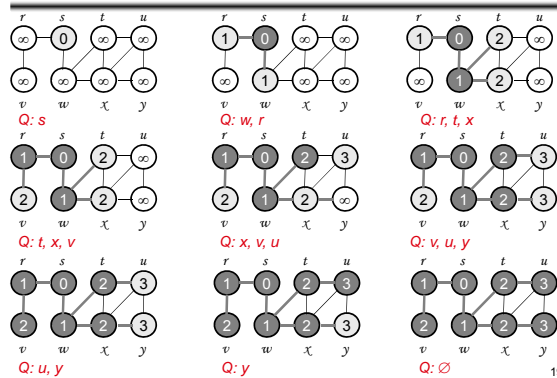
## BFS( $V, E, s$ )

- while  $Q \neq \emptyset$
- do  $u \leftarrow DEQUEUE(Q)$
- for each  $v \in Adj[u]$
- do if  $color[v] = WHITE$
- then  $color[v] = GRAY$
- $d[v] \leftarrow d[u] + 1$
- $\pi[v] = u$
- $ENQUEUE(Q, v)$
- $color[u] \leftarrow BLACK$



$Q: w, r$

## Example



19

## BFS算法分析 (Analysis of BFS)

1. for each  $u \in V - \{s\}$
2. do  $\text{color}[u] \leftarrow \text{WHITE}$
3.  $d[u] \leftarrow \infty$
4.  $\pi[u] = \text{NIL}$
5.  $\text{color}[s] \leftarrow \text{GRAY}$
6.  $d[s] \leftarrow 0$
7.  $\pi[s] = \text{NIL}$
8.  $Q \leftarrow \emptyset$
9.  $Q \leftarrow \text{ENQUEUE}(Q, s)$

$O(V)$

$\Theta(1)$

20

## Analysis of BFS

10. while  $Q \neq \emptyset$
11. do  $u \leftarrow \text{DEQUEUE}(Q)$   $\leftarrow \Theta(1)$
12. for each  $v \in \text{Adj}[u]$   $\leftarrow$  Scan  $\text{Adj}[u]$  for all vertices in the graph (搜索所有与u邻接的顶点)
13. do if  $\text{color}[v] = \text{WHITE}$
14. then  $\text{color}[v] = \text{GRAY}$
15.  $d[v] \leftarrow d[u] + 1$
16.  $\pi[v] = u$
17.  $\text{ENQUEUE}(Q, v)$   $\leftarrow \Theta(1)$
18.  $\text{color}[u] \leftarrow \text{BLACK}$

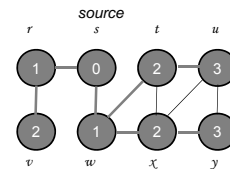
• Each vertex is scanned only once, when the vertex is dequeued (只搜索一次)  
• Sum of lengths of all adjacency lists =  $\Theta(E)$  (邻接表长度)  
• Scanning operations:  $O(E)$  (搜索操作)

- Total running time for BFS =  $O(V + E)$  运行时间

21

## 最短路径 (Shortest Paths Property)

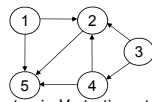
- BFS finds the shortest-path distance from the source vertex  $s \in V$  to each node in the graph (BFS确定了从源节点到其他顶点的最短路径)
- Shortest-path distance =  $\delta(s, u)$  (最短路径)
  - Minimum number of edges in any path from  $s$  to  $u$  (从s到u最小边的数目)



22

## 深度优先搜索 (Depth-First Search)

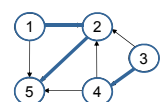
- **Input:** (输入, 图, 没有源顶点)
  - $G = (V, E)$  (No source vertex given!)
- **Goal:** (目的)
  - Explore the edges of  $G$  to "discover" every vertex in  $V$  starting at the most current visited node (从当前访问顶点开始, 探索图的边以发现图中的每个顶点)
  - Search may be repeated from multiple sources (搜索从多个源重复)
- **Output:** (输出)
  - 2 timestamps on each vertex: (每个顶点有两个时间标记)
    - $d[v]$  = discovery time (发现时间)
    - $f[v]$  = finishing time (done with examining  $v$ 's adjacency list) (结束时间, 检查v所有的邻接表)
  - Depth-first forest (深度优先森林)



23

## 深度优先搜索 (Depth-First Search)

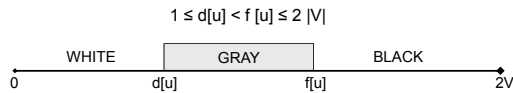
- Search "deeper" in the graph whenever possible (尽可能往深度搜索)
- Edges are explored out of the most recently discovered vertex that still has unexplored edges (最近发现的顶点v的是否有没有搜索边?)
- After all edges of  $v$  have been explored, the search "backtracks" from the parent of  $v$  (顶点v的所有边搜索完后, 回溯到v的父顶点搜索)
- The process continues until all vertices reachable from the original source have been discovered (这个过程不断继续, 直到从源顶点可以抵达的所有顶点都被发现)
- If undiscovered vertices remain, choose one of them as a new source and repeat the search from that vertex (如果图中存在还没有发现的顶点, 选择其中之一作为新的源顶点重复上述搜索过程)
- DFS creates a "depth-first forest" (DFS产生的是DFF)



24

## DFS附加数据结构 (DFS Additional Data Structures)

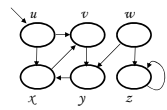
- Global variable: time-step (全局变量: 时间步骤)
  - Incremented when nodes are discovered/finished (在顶点发现和搜索完成时增加)
- color[u] – similar to BFS (颜色变量, 和BFS类似)
  - White before discovery, gray while processing and black when finished processing (白: 发现前, 灰: 发现处理中, 黑: 处理完成)
- $\pi[u]$  – predecessor of u (顶点u的父顶点)
- d[u], f[u] – discovery and finish times (发现、完成时间)



25

## DFS算法: DFS(V, E)

- for each  $u \in V$
- do color[u]  $\leftarrow$  WHITE
- $\pi[u] \leftarrow \text{NIL}$
- time  $\leftarrow 0$
- for each  $u \in V$
- do if color[u] = WHITE
- then DFS-VISIT(u)

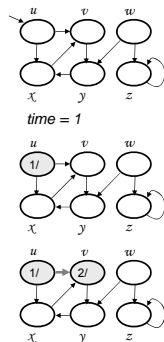


- Every time DFS-VISIT(u) is called, u becomes the root of a new tree in the depth-first forest (DFS-VISIT(u)每次被调用, u是深度优先森林中一棵新树的根节点)

26

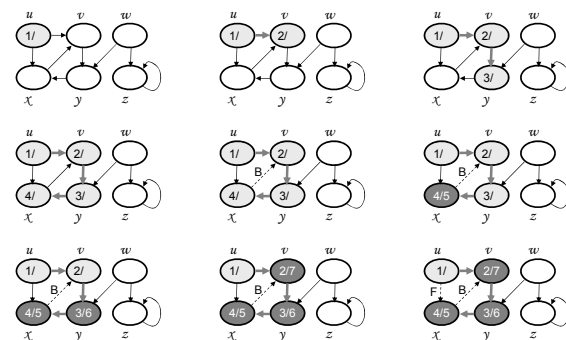
## DFS-VISIT(u)

- color[u]  $\leftarrow$  GRAY
- time  $\leftarrow$  time+1
- d[u]  $\leftarrow$  time
- for each  $v \in \text{Adj}[u]$
- do if color[v] = WHITE
- then  $\pi[v] \leftarrow u$
- DFS-VISIT(v)
- color[u]  $\leftarrow$  BLACK
- time  $\leftarrow$  time + 1
- f[u]  $\leftarrow$  time



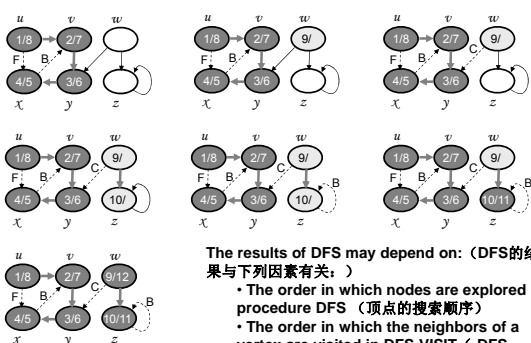
27

## Example



28

## Example (cont.)



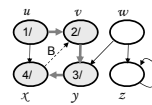
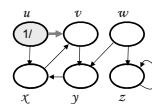
The results of DFS may depend on: (DFS的结果与下列因素有关:)

- The order in which nodes are explored in procedure DFS (顶点的搜索顺序)
- The order in which the neighbors of a vertex are visited in DFS-VISIT (DFS-VISIT中顶点邻接点的访问顺序)

29

## 边的类型(Edge Classification)

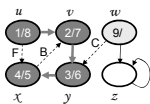
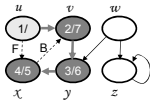
- Tree edge (reaches a WHITE vertex): (树边: 抵达白色顶点)
  - (u, v) is a tree edge if v was first discovered by exploring edge (u, v) (边(u, v)是一树边, 如果v是在探索边(u, v)时首次被发现)
- Back edge (reaches a GRAY vertex): (回边: 抵达灰色顶点)
  - (u, v), connecting a vertex u to an ancestor v in a depth first tree (边(u, v)连接顶点u至其一个深度优先树的先辈顶点v)
  - Self loops (in directed graphs) are also back edges (有向图中的自回路也是回边)



30

## Edge Classification

- Forward edge (reaches a BLACK vertex &  $d[u] < d[v]$ ): (前向边: 抵达黑色顶点, 且  $d[u] < d[v]$ )
  - Non-tree edges ( $u, v$ ) that connect a vertex  $u$  to a descendant  $v$  in a depth first tree (非树边( $u, v$ ), 连接顶点 $u$ 和它在深度优先树上的一个后裔顶点 $v$ )
- Cross edge (reaches a BLACK vertex &  $d[u] > d[v]$ ): (横跨边: 抵达黑色顶点, 且  $d[u] > d[v]$ )
  - Can go between vertices in same depth-first tree (as long as there is no ancestor / descendant relation) or between different depth-first trees (可以跨越在不同的深度优先树的顶点或同一深度优先树不同顶点间 (只要没有遗传关系))



31

## 算法分析 (Analysis of DFS( $V, E$ ))

- for each  $u \in V$
  - do  $color[u] \leftarrow WHITE$
  - $\pi[u] \leftarrow NIL$
  - time  $\leftarrow 0$
  - for each  $u \in V$
  - do if  $color[u] = WHITE$  then DFS-VISIT( $u$ )
- $\Theta(V)$
- $\Theta(V)$  – exclusive of time for DFS-VISIT (DFS-VISIT 执行时间)

32

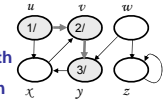
## DFS-VISIT( $u$ )分析

- $color[u] \leftarrow GRAY$
  - time  $\leftarrow time + 1$
  - $d[u] \leftarrow time$
  - for each  $v \in Adj[u]$
  - do if  $color[v] = WHITE$
  - then  $\pi[v] \leftarrow u$
  - DFS-VISIT( $v$ )
  - $color[u] \leftarrow BLACK$
  - time  $\leftarrow time + 1$
  - $f[u] \leftarrow time$
- DFS-VISIT is called exactly once for each vertex (对于每个顶点DFS-VISIT被调用一次)
- Each loop takes  $|Adj[v]|$  (循环次数)
- Total:  $\sum_{v \in V} |Adj[v]| + \Theta(V) = \Theta(V + E)$
- $\Theta(E)$

33

## DFS特性 (Properties of DFS)

- $u = \pi[v] \Leftrightarrow$  DFS-VISIT( $v$ ) was called during a search of  $u$ 's adjacency list ( $u = \pi[v] \Leftrightarrow$  DFS-VISIT( $v$ ) 在搜索 $u$ 的邻接表时被调用)
- Vertex  $v$  is a descendant of vertex  $u$  in the depth first forest  $\Leftrightarrow v$  is discovered during the time in which  $u$  is gray (深度优先树上顶点 $v$ 是顶点 $u$ 的后裔  $\Leftrightarrow$  顶点 $v$ 在顶点 $u$ 处于灰色状态时被发现)

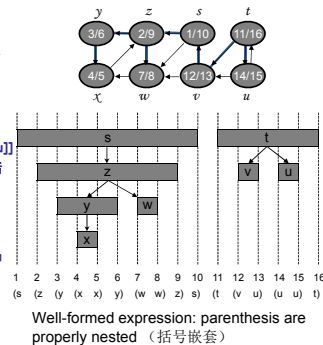


34

## 括号定理 (Parenthesis Theorem)

In any DFS of a graph  $G$ , for all  $u, v$ , exactly one of the following holds: (在图的任何DFS, 对于任何 $u, v$ , 下列情况之一成立)

- $[d[u], f[u]]$  and  $[d[v], f[v]]$  are disjoint, and neither of  $u$  and  $v$  is a descendant of the other ( $[d[u], f[u]]$  和  $[d[v], f[v]]$  不邻接,  $u$  和  $v$  互不为后裔)
- $[d[v], f[v]]$  is entirely within  $[d[u], f[u]]$  and  $v$  is a descendant of  $u$  ( $[d[u], f[u]]$  包含  $[d[v], f[v]]$ ,  $v$  是  $u$  的后裔)
- $[d[u], f[u]]$  is entirely within  $[d[v], f[v]]$  and  $u$  is a descendant of  $v$  ( $[d[v], f[v]]$  包含  $[d[u], f[u]]$ ,  $u$  是  $v$  的后裔)



35

## DFS其他特性 (Other Properties of DFS)

Corollary (推论)

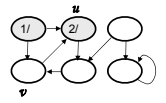
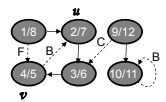
Vertex  $v$  is a proper descendant of  $u$

$\Leftrightarrow d[u] < d[v] < f[v] < f[u]$

顶点 $v$ 是顶点 $u$ 的后裔  $\Leftrightarrow d[u] < d[v] < f[v] < f[u]$

Theorem (White-path Theorem) (定理)

In a depth-first forest of a graph  $G$ , vertex  $v$  is a descendant of  $u$  if and only if at time  $d[u]$ , there is a path  $u \Rightarrow v$  consisting of only white vertices. (在图 $G$ 的任何深度优先森林中, 顶点 $v$ 是顶点 $u$ 的后裔, 当且仅当在时刻 $d[u]$ , 存在一条 $u \Rightarrow v$ 的路径只包含有白色顶点)



36

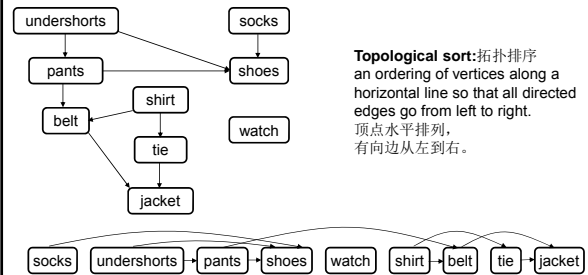
## 拓扑排序 (Topological Sort)

**Topological sort** of a directed acyclic graph  $G = (V, E)$ : a linear order of vertices such that if there exists an edge  $(u, v)$ , then  $u$  appears before  $v$  in the ordering. (拓扑排序: 找出有向无回路图  $G = (V, E)$  中顶点的一个线性序, 使得  $(u, v)$  如果是图中的一条边, 那么在这个线性序中  $u$  在  $v$  前出现)

- Directed acyclic graphs (DAGs) (有向无回路图)
    - Used to represent precedence of events or processes that have a **partial order** (表示事件的先后或有顺序的过程)
      - $a$  before  $b$  }  $a$  before  $c$      $b$  before  $c$  } What about  $a$  and  $b$ ?
      - $b$  before  $c$  }  $a$  before  $c$
- Topological sort helps us establish a **total order** (拓扑排序建立完整的顺序)

37

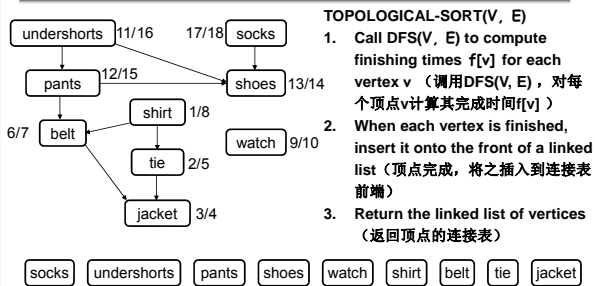
## 拓扑排序 (Topological Sort)



**Topological sort:** 拓扑排序  
an ordering of vertices along a horizontal line so that all directed edges go from left to right.  
顶点水平排列,  
有向边从左到右。

38

## 拓扑排序 (Topological Sort)



- TOPOLOGICAL-SORT( $V, E$ )**
- Call **DFS( $V, E$ )** to compute finishing times  $f[v]$  for each vertex  $v$  (调用 **DFS( $V, E$ )**, 对每个顶点  $v$  计算其完成时间  $f[v]$ )
  - When each vertex is finished, insert it onto the front of a linked list (顶点完成, 将之插入到连接表前端)
  - Return the linked list of vertices (返回顶点的连接表)

Running time:  $\Theta(V + E)$

39

## 引理 (Lemma)

A directed graph is **acyclic**  $\Leftrightarrow$  a DFS on  $G$  yields no back edges. (有向图是无回路的  $\Leftrightarrow$  深度优先搜索 DFS 不产生回边)

**Proof:**

" $\Rightarrow$ ": acyclic  $\Rightarrow$  no back edge (无回路  $\Rightarrow$  无回边)

- Assume **back edge**  $\Rightarrow$  prove **cycle** (反证法: 假设回边  $\Rightarrow$  有回路)
- Assume there is a back edge  $(u, v)$  ( $(u, v)$  是回边)
  - $\Rightarrow v$  is an ancestor of  $u$  ( $v$  是  $u$  的祖先)
  - $\Rightarrow$  there is a path from  $v$  to  $u$  in  $G$  ( $v \Rightarrow u$ ) (从  $v$  到  $u$  有通路  $v \Rightarrow u$ )
  - $\Rightarrow v \Rightarrow u +$  the back edge  $(u, v)$  yield a cycle (通路  $v \Rightarrow u$  与回边  $(u, v)$  构成回路)



40

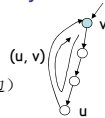
## Lemma

A directed graph is **acyclic**  $\Leftrightarrow$  a DFS on  $G$  yields no back edges.

**Proof:**

" $\Leftarrow$ ": no back edge  $\Rightarrow$  acyclic (无回边  $\Rightarrow$  无回路)

- Assume **cycle**  $\Rightarrow$  prove **back edge** (假设回路  $\Rightarrow$  回边)
  - Suppose  $G$  contains cycle  $c$  (假设图  $G$  有回路  $c$ )
  - Let  $v$  be the first vertex discovered in  $c$ , and  $(u, v)$  be the preceding edge in  $c$  (顶点  $v$  在  $c$  中最先发现,  $(u, v)$  是  $c$  中的前向边)
  - At time  $d[v]$ , vertices of  $c$  form a white path  $v \Rightarrow u$  (在  $d[v]$  时刻,  $c$  的顶点形成  $v \Rightarrow u$  白色通路)
  - $u$  is descendant of  $v$  in depth-first forest (by white-path theorem) (在深度优先森林中,  $u$  是  $v$  的后裔)
- $\Rightarrow (u, v)$  is a back edge ( $(u, v)$  是回边)

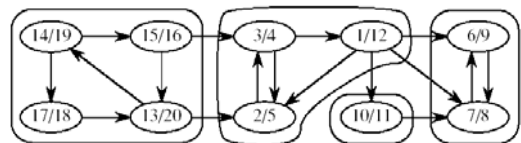


41

## 强连通分支 (Strongly Connected Components)

Given directed graph  $G = (V, E)$ :

A **strongly connected component (SCC)** of  $G$  is a maximal set of vertices  $C \subseteq V$  such that for every pair of vertices  $u, v \in C$ , we have both  $u \Rightarrow v$  and  $v \Rightarrow u$ . (有向图  $G = (V, E)$  中的强连通分支是它的顶点的极大集  $C \subseteq V$ , 在该集中, 每一对顶点  $u, v \in C$  有  $u \Rightarrow v$  和  $v \Rightarrow u$ )



42



## 图的转置(The Transpose of a Graph)

- $G^T$  = transpose of  $G$ 
  - $G^T$  is  $G$  with all edges reversed (所有边转向)
  - $G^T = (V, E^T), E^T = \{(u, v) : (v, u) \in E\}$
- If using adjacency lists: we can create  $G^T$  in  $\Theta(V + E)$  time (应用临接表, 转置时间)



43

## 确定强连通分支(Finding the SCC)

- **Observation:**  $G$  and  $G^T$  have the same SCC's ( $G^T$ 和 $G$ 有相同的SCC)
  - $u$  and  $v$  are reachable from each other in  $G \Leftrightarrow$  they are reachable from each other in  $G^T$
- Idea for computing the SCC of a DAG  $G = (V, E)$ : (有向无回路图的SCC确定方法)
  - Make two depth first searches: one on  $G$  and one on  $G^T$  (在 $G$ 和 $G^T$ 做深度搜索)



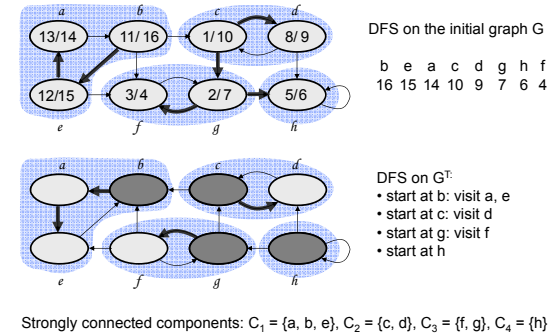
44

## STRONGLY-CONNECTED-COMPONENTS(G)

1. call DFS( $G$ ) to compute finishing times  $f[u]$  for each vertex  $u$  (在图上执行深度优先搜索, 计算每个顶点 $u$ 的完成时刻 $f[u]$ )
2. compute  $G^T$  (构成转置图 $G^T$ )
3. call DFS( $G^T$ ), but in the main loop of DFS, consider vertices in order of decreasing  $f[u]$  (as computed in first DFS) (从具有最大 $f[u]$ 的顶点开始, 在 $G^T$ 上执行深度优先搜索, 如果深度优先搜索不能到达所有的顶点, 则在余下的顶点中找一个 $f[u]$ 最大的顶点, 开始下一次深度优先搜索)
4. output the vertices in each tree of the depth-first forest formed in second DFS as a separate SCC (在最终得到的森林中的每一棵树对应一个强连通分支)

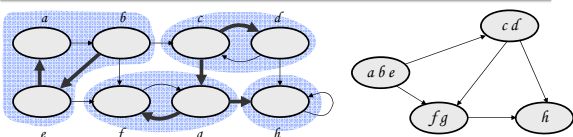
45

## Example



46

## 分支图 (Component Graph)



- The component graph  $G^{SCC} = (V^{SCC}, E^{SCC})$ : (表示方法)
  - $V^{SCC} = \{v_1, v_2, \dots, v_k\}$ , where  $v_i$  corresponds to each strongly connected component  $C_i$
  - There is an edge  $(v_i, v_j) \in E^{SCC}$  if  $G$  contains a directed edge  $(x, y)$  for some  $x \in C_i$  and  $y \in C_j$
- The component graph is a DAG (分支图是一个有向无回路图)

47

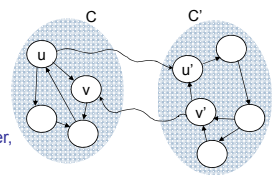
## 引理1 (Lemma 1)

Let  $C$  and  $C'$  be distinct SCC's in  $G$  ( $C$ 和 $C'$ 是图 $G$ 的SCC)  
 Let  $u, v \in C$ , and  $u', v' \in C'$   
 Suppose there is a path  $u \Rightarrow u'$  in  $G$  (假设图 $G$ 中存在通路 $u \Rightarrow u'$ )

Then there cannot also be a path  $v' \Rightarrow v$  in  $G$ .  
 (则不可能有 $v' \Rightarrow v$ )

**Proof (反证法)**

- Suppose there is a path  $v' \Rightarrow v$  (假设存在 $v' \Rightarrow v$ )
- There exists  $u \Rightarrow u' \Rightarrow v'$
- There exists  $v' \Rightarrow v \Rightarrow u$
- $u$  and  $v'$  are reachable from each other, so they are not in separate SCC's: contradiction! (和双向可连通, 与他们在不同的SCC相矛盾)

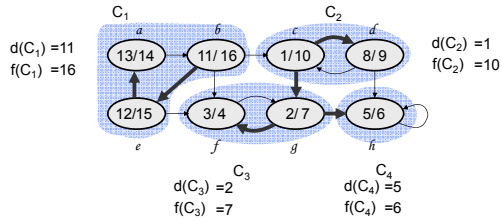


48



## 符号标记 (Notations)

- Extend notation for  $d$  (starting time) and  $f$  (finishing time) to sets of vertices  $U \subseteq V$ :
  - $d(U) = \min_{u \in U} \{d[u]\}$  (earliest discovery time) (最早发现时间)
  - $f(U) = \max_{u \in U} \{f[u]\}$  (latest finishing time) (最晚结束时间) (用于SCC)

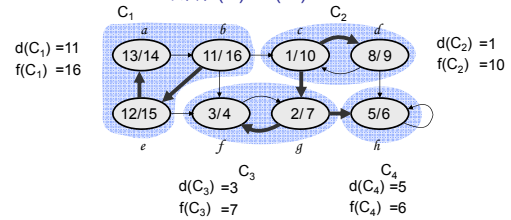


49

## 引理2 (Lemma 2)

- Let  $C$  and  $C'$  be distinct SCCs in a directed graph  $G = (V, E)$ . If there is an edge  $(u, v) \in E$ , where  $u \in C$  and  $v \in C'$  then  $f(C) > f(C')$ .

( $C$ 和 $C'$ 是有向图 $G = (V, E)$ 的SCC, 如果存在边 $(u, v) \in E$ ,  $u \in C$ ,  $v \in C'$ , 则有 $f(C) > f(C')$ 。)



50