数据结构和算法实习

郭 炜

学会程序和算法，走遍天下都不怕!

讲义照片均为郭炜拍摄

图的连通性相关问题

拓扑排序

有向无环图的拓扑排序

对一个有向无环图(Directed Acyclic Graph简称DAG)G进行拓扑排序，是将G中所有顶点排成一个线性序列，使得图中任意一对顶点u和v，若边(u,v)∈E(G)，则u在线性序列中出现在v之前。

有向无环图的拓扑排序

拓扑排序做法：

1) 找出所有入度为0的点放入队列
2) 从队头拿出一个节点p输出，删除所有p连出的边（p,x）。若发现删边后x的入度变为0，则将x加入队列
3) 若队列不为空，则转2)，否则结束

有向图强连通分支

瑞士布里茨恩湖

有向图强连通分支的定义

- 在有向图G中，如果任意两个不同的顶点相互可达，则称该有向图是强连通的。有向图G的极大强连通子图称为G的强连通分支。

有向图强连通分支的Tarjan算法

- 用DFS的方法遍历有向图（每次任选没访问过的点作为起点），用dfn[i]表示编号为i的节点在整个DFS过程中的访问序号(也可以叫做开始时间）。在DFS过程中会形成一棵或若干棵搜索树。在一棵搜索树上越先遍历到的节点，显然dfn的值就越小。dfn值越小的节点，就称为越“早”。

- 用low[i]表示从i节点出发DFS过程中i下方节点(开始时间大于dfn[i]，且由i可达的节点）所能到达的最早的，在当前搜索路径上的节点的开始时间。初始时low[i]=dfn[i]

有向图强连通分支的Tarjan算法

- DFS过程中，碰到哪个节点，就将哪个节点入栈。栈中节点只有在其所属的强连通分量已经全部求出时，才会出栈。

- 如果发现某节点u有边连到当前搜索路径上的节点v，则更新u的low 值为min(low[u],dfn[v])，若low[u]被更新为dfn[v],则表明目前发现u可达的最早的节点是v.

有向图强连通分支的Tarjan算法

- 对于u的子节点v，从v出发进行的DFS结束回到u时，使得 low[u] = min(low[u],low[v])。因为u可达v,所以v可达的最早的节点，也是u可达的。

- 如果一个节点u，从其出发进行的DFS已经全部完成并回到u，而且此时其low值等于dfn值，则说明u可达的所有节点，都不能到达任何比u早的节点 ---- 那么该节点u就是一个强连通分量在DFS搜索树中的根。

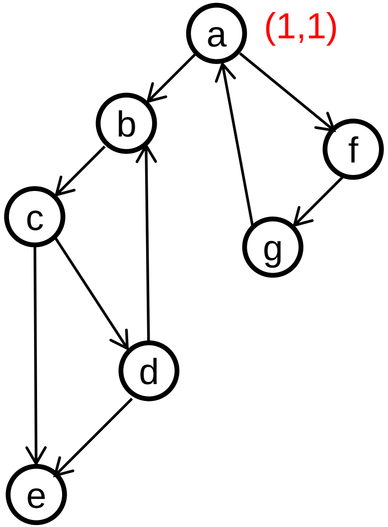
- 此时，显然栈中u上方的节点，都是不能到达比u早的节点的。将栈中节点弹出，一直弹到u(包括u),弹出的节点就构成了一个强连通分量.

有向图强连通分支的Tarjan算法

```
void Tarjan(u) {
  dfn[u]=low[u]= ++index //index是开始时间
  stack.push(u)
  for each (u, v) in E { // E是边集合
        if (v is not visited) {
                tarjan(v)
                low[u] = min(low[u], low[v])
        }
        else if (v in stack) {
                low[u] = min(low[u], dfn[v])
        }
  }
  if (dfn[u] == low[u]) { //u是一个强连通分量的根
        repeat
                v = stack.pop
                print v
        until (u== v)
  } //退栈，把整个强连通分量都弹出来
} //复杂度是O(E+V)的
```

(dfn,low)

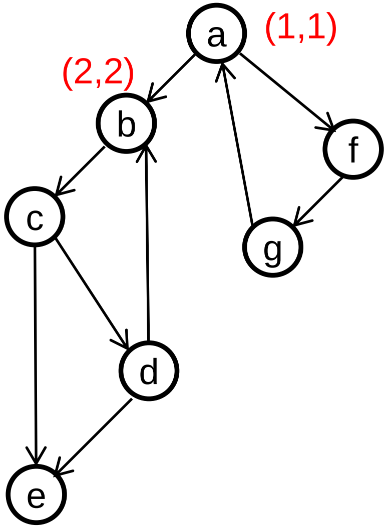
栈

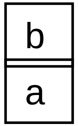


a

(dfn,low)

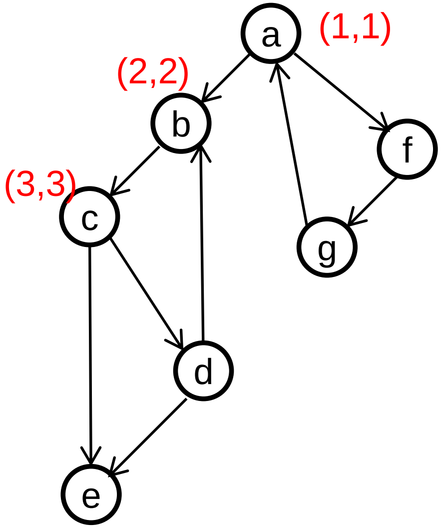
栈

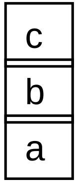




(dfn,low)

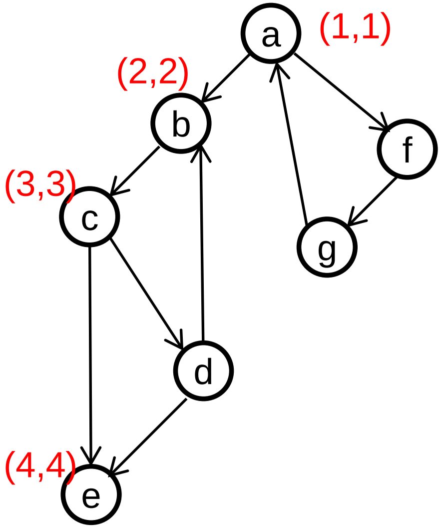
栈

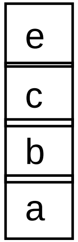




(dfn,low)

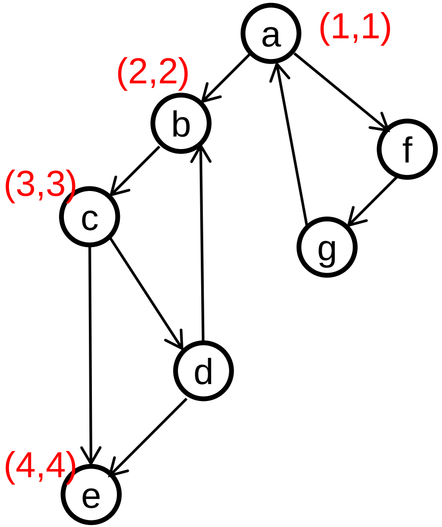
栈

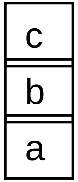




(dfn,low)

栈



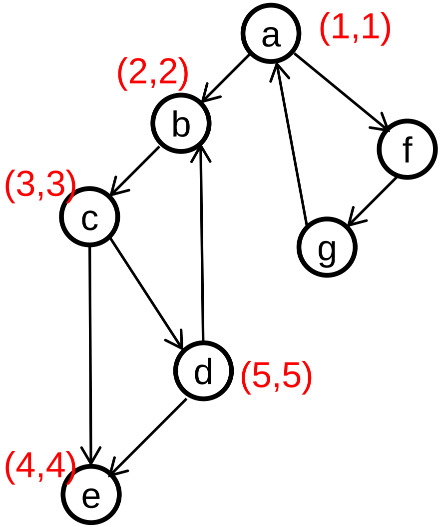


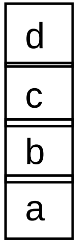
强连通分量：

{e}

(dfn,low)

栈



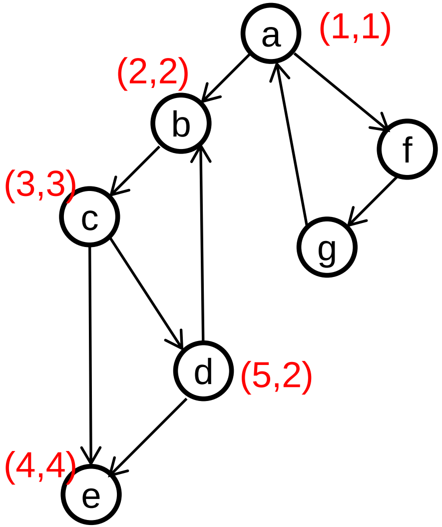


强连通分量:

{e}

(dfn,low)

栈



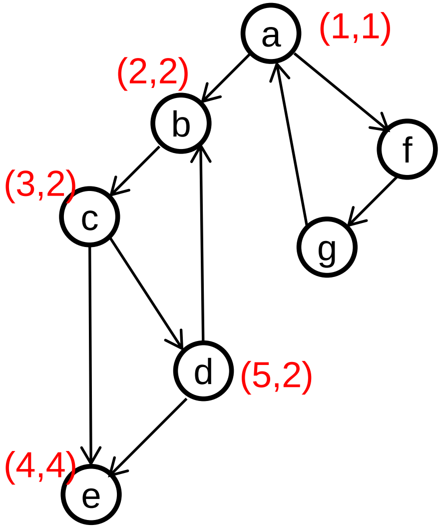
d
c
b
a

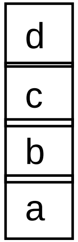
强连通分量：

{e}

(dfn,low)

栈



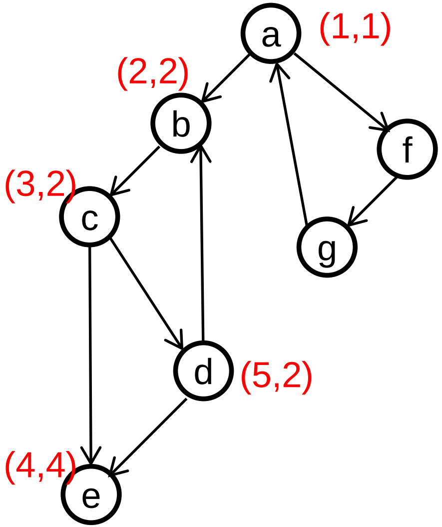


强连通分量：

{e}

(dfn,low)

栈



a

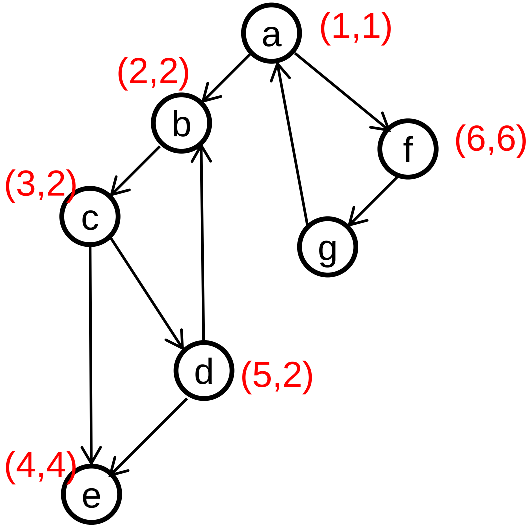
强连通分量：

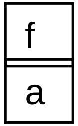
{e}

{b c d}

(dfn,low)

栈





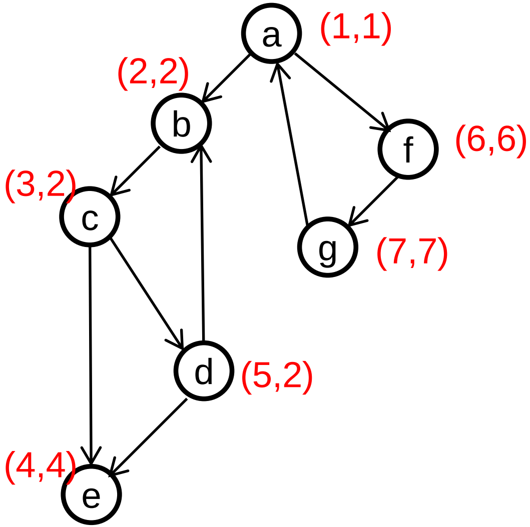
强连通分量：

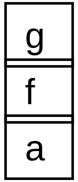
{e}

{b c d}

(dfn,low)

栈



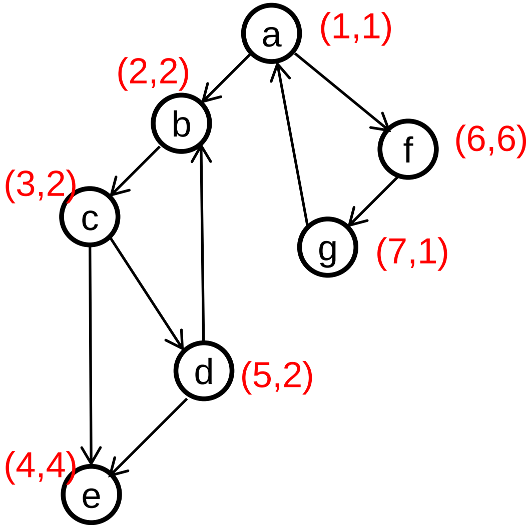


强连通分量：

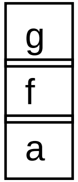
{e}

{b c d}

(dfn,low)



栈



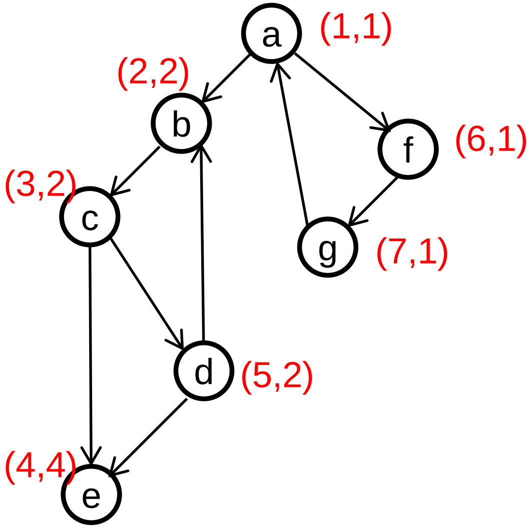
强连通分量:

{e}

{b c d}

(dfn,low)

栈



强连通分量：

{e}

{b c d}

{a f g}

有向图强连通分支的Tarjan算法

- 为什么从u出发的DFS全部结束回到u时，若dfn[u]=low[u], 此时将栈中u及其上方的节点弹出，就找到了一个强连通分量?

此时所有节点分成以下几类:

1） 还没被访问过的节点 —— 从u不可达
2) 栈中比u早的节点(栈中在u下方) —— 可达u，但从u不可达
因为low[u]=dfn[u]
3) 栈中比u晚的节点(栈中在u上方) —— 和u互相可达
4) 栈中的u —— 和u互相可达
5) 曾经在u之前入栈（访问过），又出了栈的节点 —— 不可达u
否则应该还在栈里，u下面
6) 曾经在u之后入栈（访问过），又出了栈的节点 —— 从u可达，但不可达u

有向图强连通分支的Tarjan算法

证：3）栈中比u晚的节点x(栈中在u上方) 和u互相可达

由u可达显然。

则寻找x所能到达的最早的，栈里面的节点y:

1) y比u早或y就是u: y可达u → x 也可达u
2) y比u晚: 不可能，因为:

　　a) low[y] < dfn[y]: 不可能。low[y] < dfn[y] 说明y可达比y更早的节点，则x也可达比y更早的节点。这和y是x可达的最早的节点矛盾。

　　b) low[y] = dfn[y]: 也不可能。若为真，则由y出发做DFS回到y时，栈中y及其上方节点应该已经被弹出栈了，y上方的x当然也已经不再栈中，这和x是第3类节点矛盾。

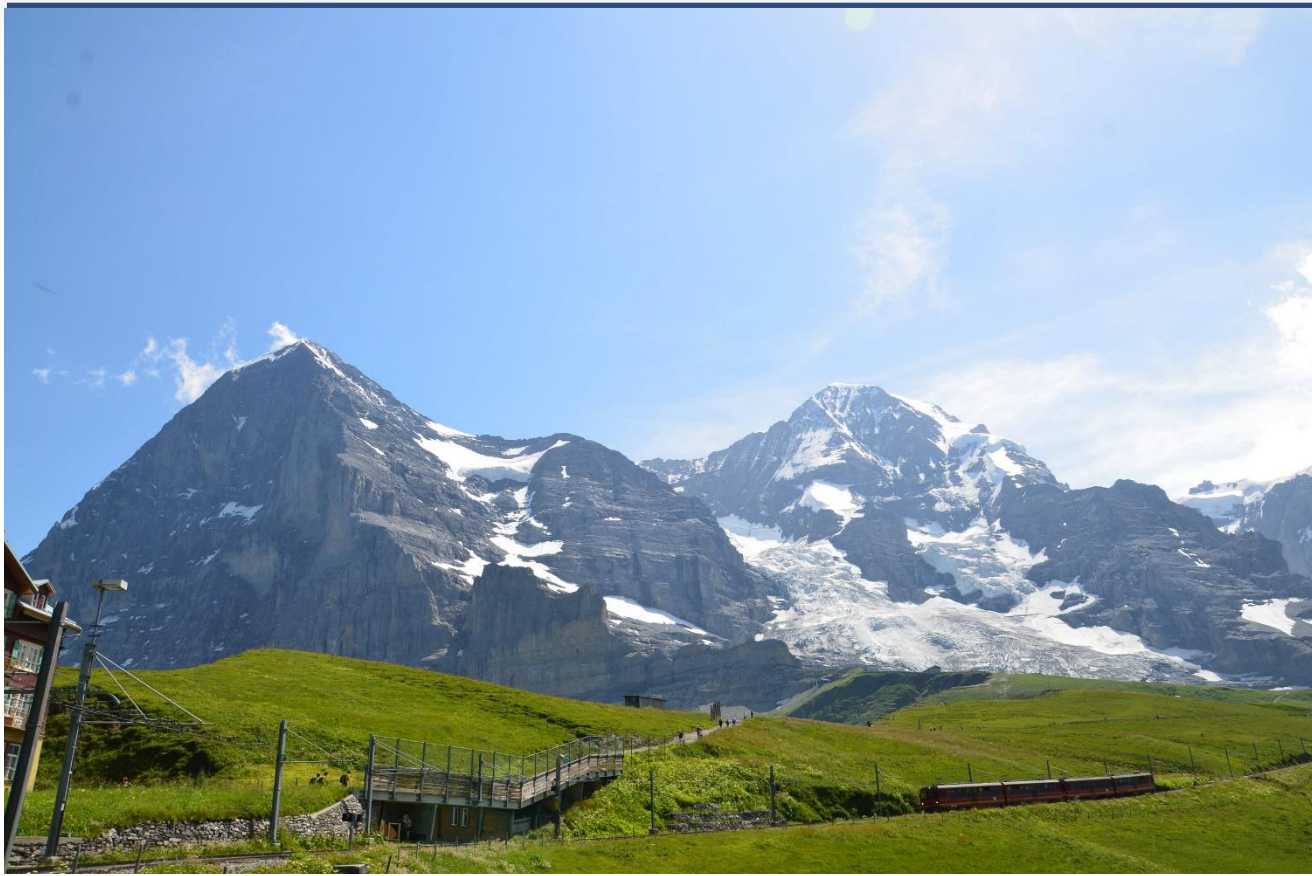
有向图强连通分支的Tarjan算法

➢ 证： 6）曾经在u之后入栈（访问过），又出了栈的节点 —— 从u可达，但不可达u

任取此类节点x。x之所以已经被弹出栈，定是因为以下2种原因之一：

1) low[x] = dfn[x]
2) 在栈里，x位于某个y节点上方(由y可达) ,且y满足条件:最终的 low[y] = dfn[y]。因为y曾经出现在u的上方，所以y一定晚于u。因为low[x]不可能小于等于dfn[u](否则low[y]就也会小于等于dfn[u],这和low[y]=dfn[y]矛盾），所以x到达不了u及比u早的节点。

例题1
Popular Cows



瑞士少女峰

例题1： POJ2186 Popular Cows

- 有N头牛。如果a喜欢b，b喜欢c，则a也会喜欢c。告诉你M个喜欢关系，比如(a,b)表示a喜欢b。问有多少头牛是被所有牛都喜欢的。

  N<= 10,000，M<= 50,000

例题1： POJ2186 Popular Cows

- 有N头牛。如果a喜欢b，b喜欢c，则a也会喜欢c。告诉你M个喜欢关系，比如(a,b)表示a喜欢b。问有多少头牛是被所有牛都喜欢的。
  N<= 10,000，M<= 50,000

- 本质：给定一个有向图，求有多少个顶点是由任何顶点出发都可达的。

例题1： POJ2186 Popular Cows

- 有用的定理:

有向无环图中唯一出度为0的点，一定可以由任何点出发均可达（由于无环，所以从任何点出发往前走，必然终止于一个出度为0的点）

例题1： POJ2186 Popular Cows

1. 求出所有强连通分量

2. 每个强连通分量缩成一点，则形成一个有向无环图DAG。

3. DAG上面如果有唯一的出度为0的点，则该点能被所有的点可达。那么该点所代表的连通分量上的所有的原图中的点，都能被原图中的所有点可达，则该连通分量的点数，就是答案。

4. DAG上面如果有不止一个出度为0的点，则这些点互相不可达，原问题无解，答案为0

例题1： POJ2186 Popular Cows

- 缩点构造新图：把不同强连通分量的点染不同颜色，有几种颜色，新图就有几个点。扫一遍老图所有的边，跨两种颜色的边，加到新图上（注意不要加了重边）。

  新图邻接表： `vector< set<int> > G(colorNum);`

  set便于去重

- 可以不构造新图，只要把不同强连通分量的点染不同颜色，然后考察各种颜色的点有没有连到别的颜色的边即可(即其对应的缩点后的DAG图上的点是否有出边）。

例题2
Network of Schools

瑞士少女峰

例题2： POJ1236: Network of Schools

- N(2<N<100)个学校之间有单向的网络，每个学校得到一套软件后，可以通过单向网络向周边的学校传输。

- 问题1）开始至少需要向多少个学校发放软件，才能使网络内所有的学校最终都能得到软件

- 问题2) 至少需要添加几条传输线路(边)，使任意向一个学校发放软件后，经过若干次传送，网络内所有的学校最终都能得到软件。

例题2： POJ1236: Network of Schools

- 给定一个有向图，求：

1) 至少要选几个顶点，才能做到从这些顶点出发，可以到达全部顶点

2) 至少要加多少条边，才能使得从任何一个顶点出发，都能到达全部顶点

有用的定理:

有向无环图中所有入度不为0的点，一定可以由某个入度为0的点出发可达。(由于无环，所以从任何入度不为0的点往回走，必然终止于一个入度为0的点)

例题2： POJ1236: Network of Schools

- 1. 求出所有强连通分量
- 2. 每个强连通分量缩成一点，则形成一个有向无环图DAG。
- 3. DAG上面有多少个入度为0的顶点，问题1的答案就是多少

例题2： POJ1236: Network of Schools

- 问题2）在DAG上要加几条边，才能使得DAG变成强连通的?
- 加边的方法：
- 要为每个入度为0的点添加入边，为每个出度为0的点添加出边
- 假定有 n 个入度为0的点，m个出度为0的点，max(m, n)就是第二个问题的解(证明略)

无向连通图求割点和桥

德国新天鹅堡

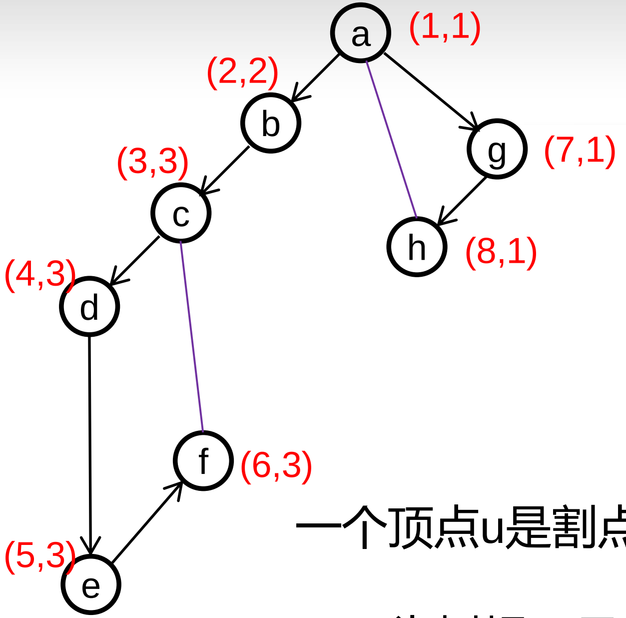
无向连通图求割点和桥

- 无向连通图中，如果删除某点后，图变成不连通，则称该点为割点。
- 无向连通图中，如果删除某边后，图变成不连通，则称该边为桥。

无向连通图求桥和割点的Tarjan算法

思路和有向图求强连通分量类似

深度优先遍历图形成一棵搜索树, dfn[u]定义和前面类似, low[u]定义为u或者u的子树中能够通过非父子边(父子边就是搜索树上的边）到达的最早的节点的DFS开始时间



割点：

c b a

桥

ab
bc

一个顶点u是割点，当且仅当满足(1)或(2)

(1) u为树根，且u有多于一个子树。
(2) u不为树根，且存在(u,v)为树枝边(或称父子边，即u为v在搜索树中的父亲)，使得dfn(u)<=low(v)。

带箭头的边是树枝边
紫色边是反向边
非树枝边不可能是桥

一条边(u,v)是桥，当且仅当(u,v)为树枝边，且满足dfn(u)<low(v)（前提是其没有重边）。

无向连通图求桥和割点的Tarjan算法

```
Tarjan(u) {
    dfn[u]=low[u]=++index
    for each (u, v) in E {
        if (v is not visited)
            Tarjan(v)
            low[u] = min(low[u], low[v])
             dfn[u]<low[v] ⇔  (u, v) 是桥
        }
        else {
            if(v不是u 的父节点)
                low[u] = min(low[u], dfn[v])
        }
    }
    if (u is root)
        u 是割点 <=> u在搜索树上至少两个子节点
    else
        u 是割点 <=> u 有一个子节点v, 满足dfn[u]<= low[v]
}
```

无向连通图求桥和割点的Tarjan算法

- 也可以先用Tajan()进行dfs算出所有点的low和dfn值，并记录dfs过程中每个点的父节点，然后再把所有点看一遍，看其low和dfn,以找出割点和桥。

- 找桥的时候，要注意看有没有重边。有重边，则不是桥。

无重边连通无向图求割点和桥

Input: （11点13边）

11 13
1 2
1 4
1 5
1 6
2 11
2 3
4 3
4 9
5 8
5 7
6 7
7 10
11 3

output:
1
4
5
7
5,8
4,9
7,10

```
//无重边连通无向图求割点和桥的程序
#include <iostream>
#include <vector>
using namespace std;

#define MyMax 200
typedef vector<int> Edge;
vector<Edge> G(MyMax);
bool Visited[MyMax] ;
int dfn[MyMax] ;
int low[MyMax] ;
int Father[MyMax]; //DFS树中每个点的父节点
bool bIsCutVetext[MyMax]; //每个点是不是割点
int nTime; //Dfs时间戳
int n,m; //n是点数，m是边数
```

```
void Tarjan(int u, int father) //father 是u的父节点
{
      Father[u] = father;
      int i,j,k;
      low[u] = dfn[u] = nTime ++;
      for( i = 0;i < G[u].size() ;i ++ ) {
            int v = G[u][i];
            if( ! dfn[v]) {
                  Tarjan(v,u);
                  low[u] = min(low[u],low[v]);
            }
            else if( father != v ) //连到父节点的回边不考虑，否则求不出桥
                  low[u] = min(low[u],dfn[v]);
      }
}
```

```
void Count()
{ //计算割点和桥
      int i,nRootSons = 0;
      Tarjan(1,0);
      for( i = 2;i <= n;i ++ ) {
            int v = Father[i];
            if( v == 1 )
                  nRootSons ++; //DFS树中根节点有几个子树
            else if( dfn[v] <= low[i])
                  bIsCutVetext[v] = true;
      }
      if( nRootSons > 1)
            bIsCutVetext[1] = true;
      for( i = 1;i <= n;i ++ )
            if( bIsCutVetext[i] )
                  cout << i << endl;
      for( i = 1;i <= n;i ++) {
            int v = Father[i];
            if(v >0 &&  dfn[v] < low[i])
                  cout << v << "," << i <<endl;
      }
}
```

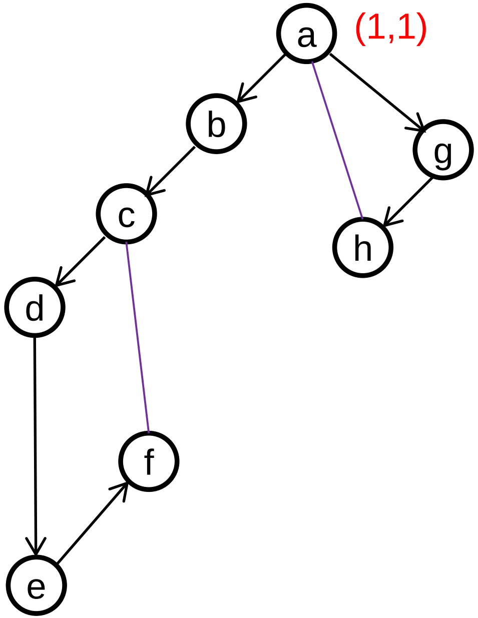
```
int main()
{
	int u,v;
	int i;

	nTime = 1;
	cin >> n >> m ; //n是点数，m是边数
	for(  i = 1;i <= m;i ++ ) {
		cin >> u >> v; //点编号从1开始
		G[v].push_back(u);
		G[u].push_back(v);
	}
	memset( dfn,0,sizeof(dfn));
	memset( Father,0,sizeof(Father));
	memset( bIsCutVetext,0,sizeof(bIsCutVetext));
	Count();
	return 0;
}
```

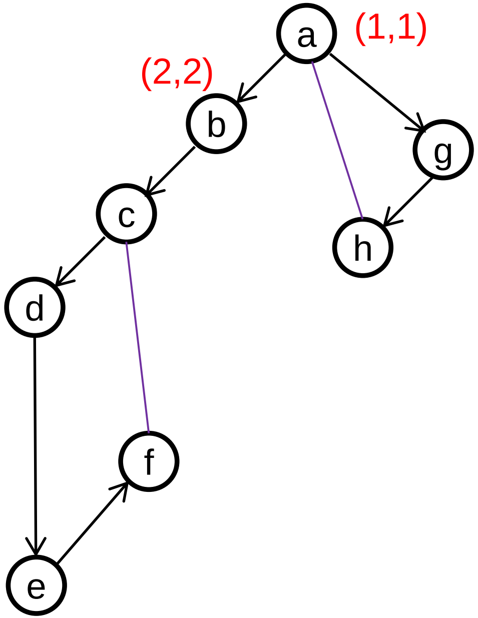
求无向连通图点双连通分支

瑞士苏黎世湖

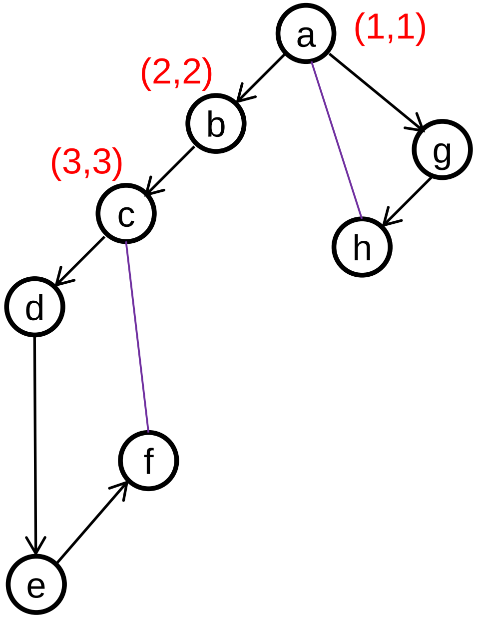
求无向连通图点双连通分支
(不包含割点的极大连通子图)

- 求割点的过程中就能顺便把每个点双连通分支求出
- 建立一个栈，存储当前双连通分支，在搜索图时，每找到一条树枝边或反向边(连到树中祖先的边），就把这条边加入栈中。如果遇到某树枝边（u,v) 满足dfn(u)<=low(v)，说明u是一个点双连通分量的根，此时把边从栈顶一个个取出，直到遇到了边(u,v)，取出的这些边与其关联的点，组成一个点双连通分支。割点可以属于多个点双连通分支，其余点和每条边只属于且属于一个点双连通分支。


a
(1,1)
b
g
c
h
d
f
e

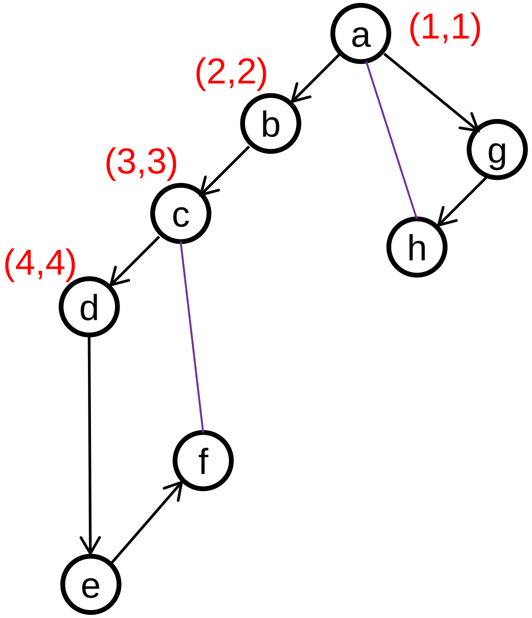


ab

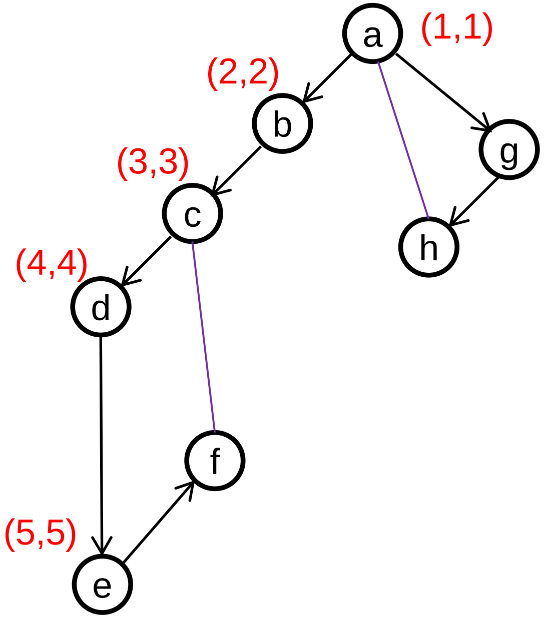


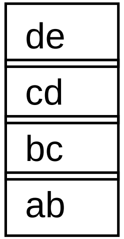
bc

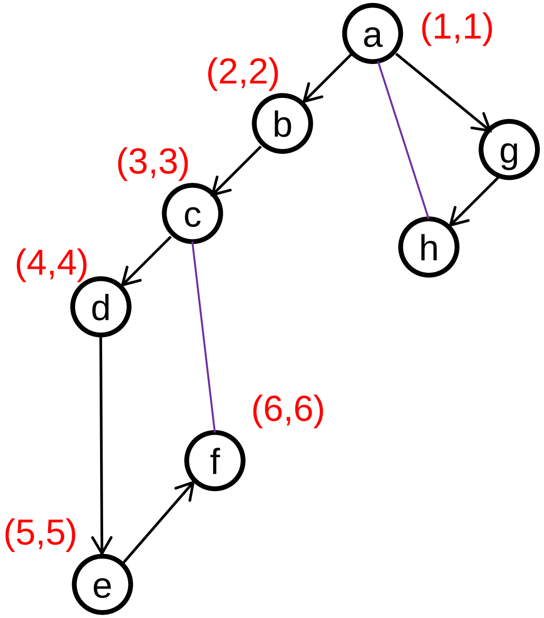
ab



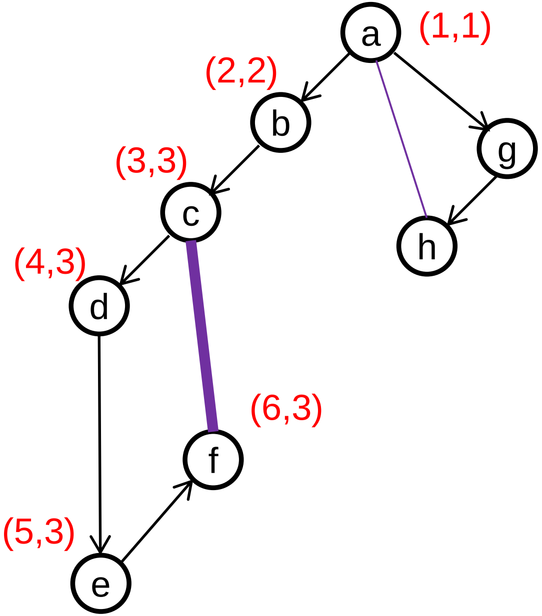
cd
bc
ab


a
(1,1)
(2,2)
b
g
(3,3)
c
h
(4,4)
d
f
(5,5)
e


de
cd
bc
ab

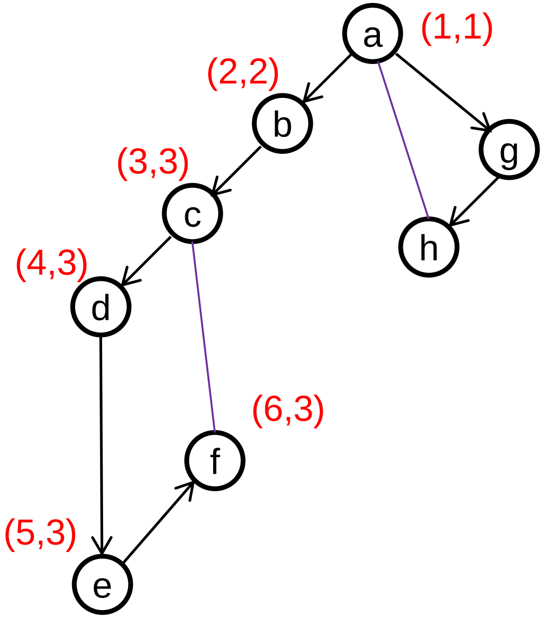


ef
de
cd
bc
ab



fc
ef
de
cd
bc
ab

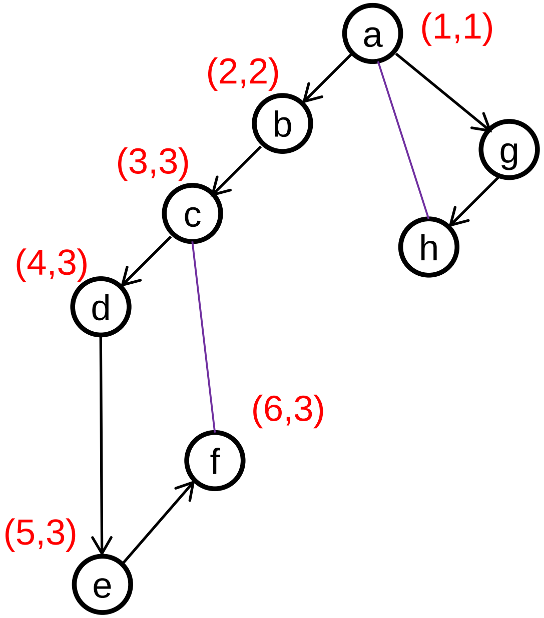
不要让 fc 入栈两次!



点双连通分量：

{ fc,ef,de,cd }

bc
ab

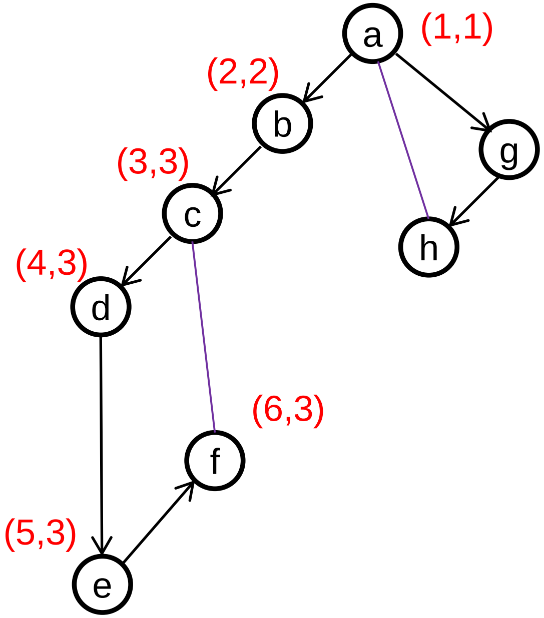


点双连通分量：

{ fc,ef,de,cd }

{bc }

ab

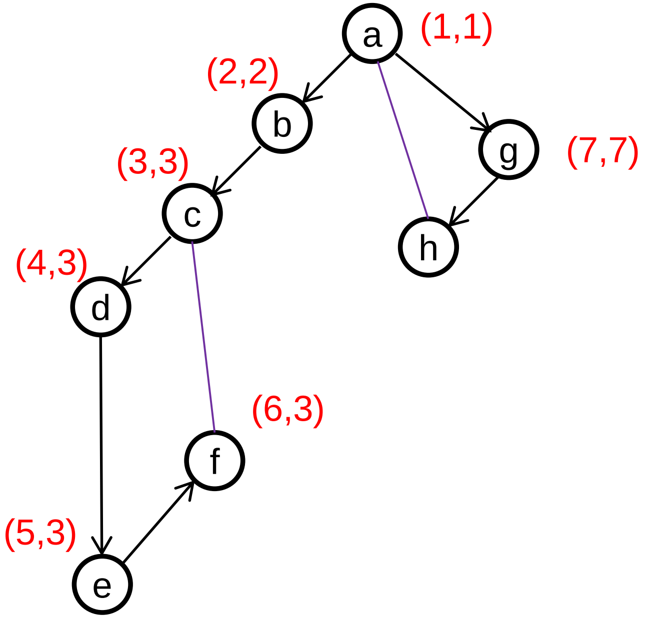


点双连通分量：

{ fc,ef,de,cd }

{ bc }

{ab}



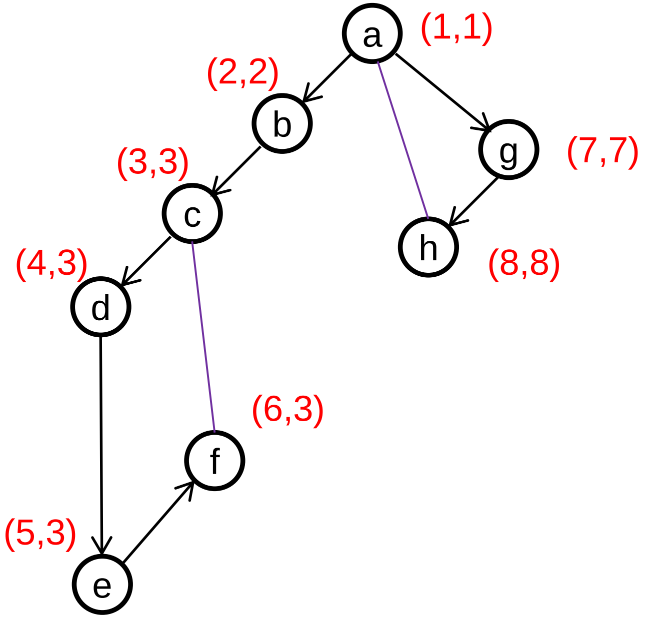
点双连通分量：

{ fc,ef,de,cd }

{ bc }

{ab}

ag



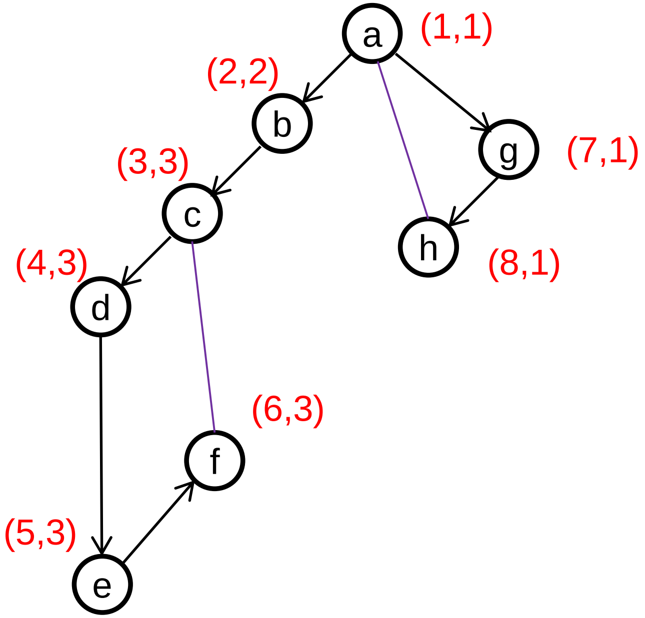
点双连通分量：

{ fc,ef,de,cd }

{ bc }

{ab}

gh
ag



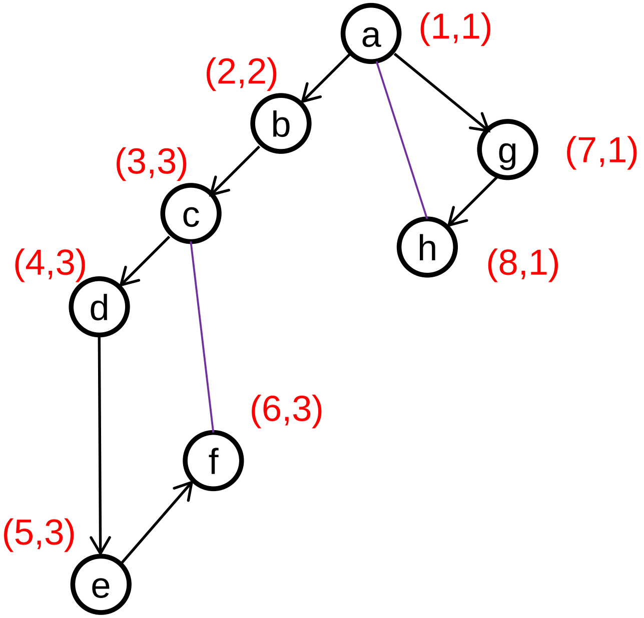
ha
gh
ag

点双连通分量：

{ fc,ef,de,cd }

{ bc }

{ab}



点双连通分量：

{ fc,ef,de,cd }

{ bc }

{ab}

{ha,gh,ag}

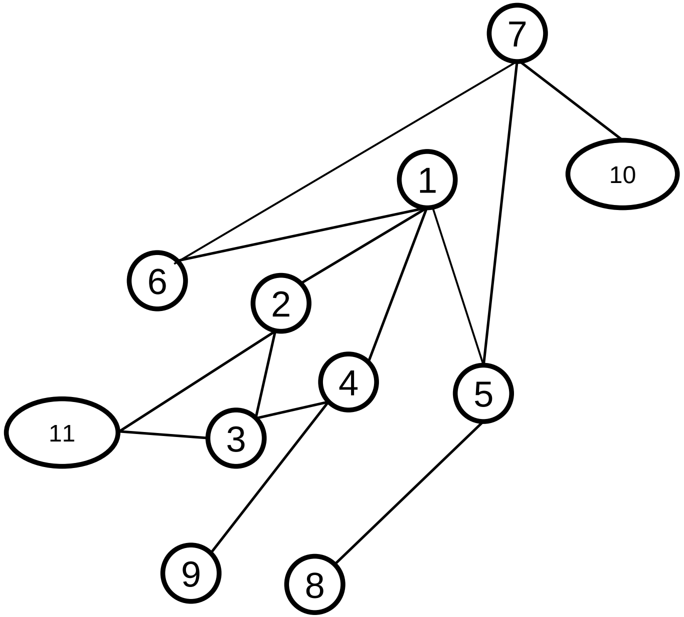
求无向连通图点双连通分量（没有割点的连通分量），假定没有重边

Input: （11点13边）

11 13
1 2
1 4
1 5
1 6
2 11
2 3
4 3
4 9
5 8
5 7
6 7
7 10
11 3

output:

Block No: 1
4,9
Block No: 2
4,1
3,4
3,2
11,3
2,11
1,2
Block No: 3
5,8
Block No: 4
7,10
Block No: 5
6,1
7,6
5,7
1,5


7
10
1
6
2
4
5
11
3
9
8

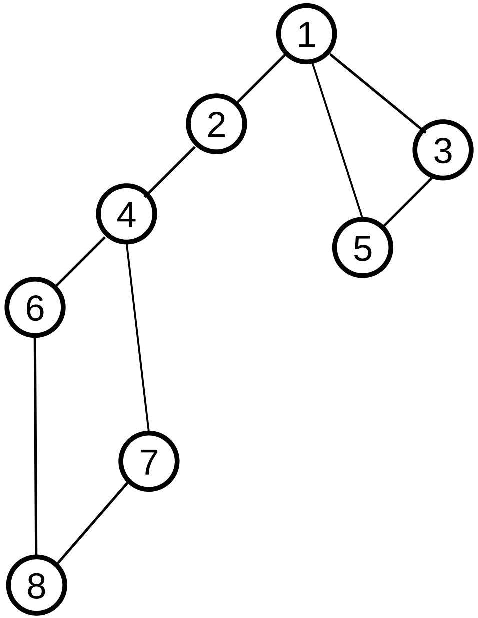
求无向连通图点双连通分量（没有割点的连通分量），假定没有重边

Input: （8点9边）

8 9
1 2
1 3
1 5
3 5
2 4
4 6
4 7
6 8
7 8

output:

Block No: 1
7,4
8,7
6,8
4,6
Block No: 2
2,4
Block No: 3
1,2
Block No: 4
5,1
3,5
1,3


1
2
3
4
5
6
7
8

```
//求无向连通图点双连通分量（没有割点的连通分量），假定没有重边
#include <iostream>
#include <vector>
#include <queue>
using namespace std;
#define MyMax 200
typedef vector<int> Edge;
vector<Edge> G(MyMax);
int dfn[MyMax] ;
int low[MyMax] ;
int nTime;
int n,m;  //n是点数，m是边数
struct Edge2
{
      int u;
      int v;
      Edge2(int u_,int v_):u(u_),v(v_) { }
};
deque<Edge2> Edges;  //栈
int nBlockNo = 0; //点双连通分量个数
```

```
void Tarjan(int u, int father)
{
        int i,j,k;
        low[u] = dfn[u] = nTime ++;
        for( i = 0;i < G[u].size() ;i ++ ) {
                int v = G[u][i];
                if( ! dfn[v]) { //v没有访问过
                        //树边要入栈
                        Edges.push_back(Edge2(u,v));
                        Tarjan(v,u);
                        low[u] = min(low[u],low[v]);
                        Edge2 tmp(0,0);
                        if(dfn[u] <= low[v]) {
                        //从一条边往下走，走完后发现自己是割点，则栈中的边一定全是和自
己在一个双连通分量里面
                    //根节点总是和其下的某些点在同一个双连通分量里面
                                cout << "Block No: " << ++ nBlockNo
                               << endl;
```

```
                    do {
                            tmp = Edges.back();
                            Edges.pop_back ();
                            cout << tmp.u << "," <<
                                    tmp.v << endl;
                    }while ( !(tmp.u == u &&
                                   tmp.v == v) );
                }
        }  // 对应if( ! dfn[v]) {
        else {
                if( v != father ) {//u连到父节点的回边不考虑
                        low[u] = min(low[u],dfn[v]);
                        if( dfn[u] > dfn[v])
                            //连接到祖先的回边要入栈，但是连接到儿子的边，此处
肯定已经入过栈了，不能再入栈
                            Edges.push_back(Edge2(u,v));
                }
        }
    } //对应    for( i = 0;i < G[u].size() ;i ++ ) {
}
```

```
int main()
{

	int u,v;
	int i;
	nTime = 1;
	cin >> n >> m ;   //n是点数，m是边数
	nBlockNo = 0;
	for(  i = 1;i <= m;i ++ ) {
		cin >> u >> v; //点编号从1开始
		G[v].push_back(u);
		G[u].push_back(v);
	}
	memset( dfn,0,sizeof(dfn));
	Tarjan(1,0);
	return 0;
}
```

求无向连通图边双连通分支

奥地利维也纳美泉宫

求无向连通图边双连通分支

边双连通分支：不包含桥的极大连通子图

只需在求出所有的桥以后，把桥边删除，原图变成了多个连通块，则每个连通块就是一个边双连通分支。桥不属于任何一个边双连通分支，其余的边和每个顶点都属于且只属于一个边双连通分支。

例题：POJ 3352 Road Construction

- 给定一个图，要求加入最少的边，使得最后得到的图为一个边双连通分支

- 可以求出所有的桥，把桥删掉。然后把所有的连通分支求出来，这些连通分支就是原图中的双连通分支。把它们缩成点，然后添上刚才删去的桥，就构成了一棵树。在树上添边使得树变成一个双连通分支即可。

例题：POJ 3352 Road Construction

- 需要添加的边数，就是（叶子节点数 + 1) /2 个 (去尾取整)

命题：一棵有n(n>=2)个叶子结点的树，至少(只需）要添加ceil(n/2)条边，才(就)能转变为一个没有桥的图。或者说，使得图中每条边，都至少在一个环上。

证明:

这里只证明n为偶数的情况。n为奇数的证明类似。

先证明添加n/2条边一定可以达成目标。

n=2时，显然只需将这两个叶子间连一条边即可。命题成立。
设n=2k(k>=1)时命题成立，即AddNum(2k)=k。下面将推出n=2(k+1)时命题亦成立

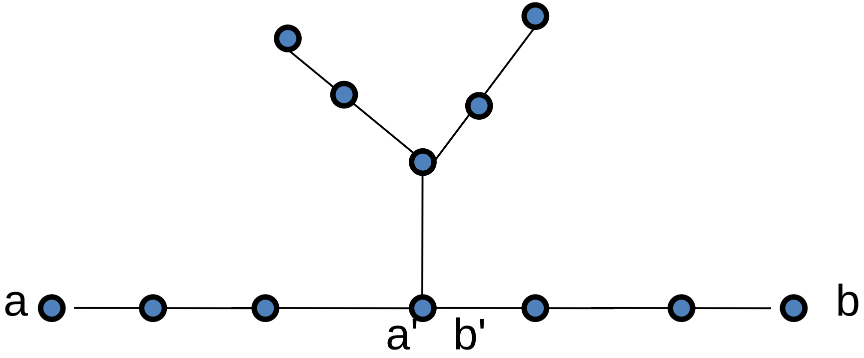
n=2k+2时，选取树中一条迹(无重复点的路径）， 设其端点为a,b；并设离a最近的度>=3的点为a',同理设b'。
（关于a‘和b’的存在性问题：由于a和b的度都为1,因此树中其它的树枝必然从迹<a,b>之间的某些点引出。否则整棵树就是迹<a,b>，n=2<2k+2，不可能。）

a’ b’不重合时:

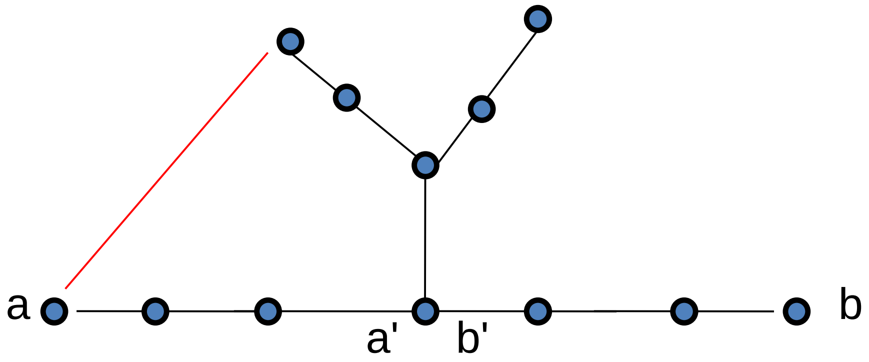
在a,b间添一条边，则迹<a,b>上的所有边都已不再是桥。这时，将刚才添加的边，以及aa‘之间，bb’之间的边都删去，得到一棵新的树。

由于a‘和b’的度>=3，所以删除操作不会使他们变成叶子。因此新的树必然比原树少了两个叶子a,b，共有2k个叶子。由归纳知需要再加k条边。因此对n=2k+2的树，一共要添加k+1条边。

a’ b’ 重合时：



a’ b’ 重合时：



将a和一个非b的叶子节点x连上，然后将环缩点至 a’。
因为叶子节点是偶数，所以必然还存在一个非b非x的叶子节点不在环上， 因此a’不会变成叶子节点，于是新图比原图少2个叶子节点。

再证明n/2是最小的解。

只有一个叶子结点被新加的边覆盖到，才有可能使与它相接的那条边进入一个环中。而一次加边至多覆盖2个叶子。因此n个叶子至少要加n/2条边。

POJ相关题目

- 1236，3180，2762（强连通+拓扑排序），2553，3114（强连通+dijkstra)，3160(强连通+DP）