离 散 数 学 2

《图 论》

北航计算机学院

第六章 树、有向树和有序数

主要内容

1. 图论的基本概念
2. 子图和图的运算
3. 路径、回路和连通图
4. 欧拉图和哈密顿图
5. 图的矩阵表示
6. 树、有向树和有序树

树、有向树、有序树

目的：树的六种定义，了解分支、森林、生成树、生成森林、最小生成树、枝、弦、基本回路、有向树、有向森林、二叉树（完全二元树）、最优二叉树、有序树、有序森林、定位二元有序树等概念和性质；掌握求最小生成树、最优二叉树的算法、定位二元有序树和有序森林的双射关系，以及有关的证明方法；

重点：树的六种定义，各种概念、算法及基本的证明思路；

难点：通过树的六种定义方式如何发现树的各种性质，大量相关知识点在证明种的综合运用。

D. 课堂思考

概念图谱

无向图G

什么是树？ 树 $\xrightarrow{G\text{去掉连通约束}}$ 森林

7个条件(连通;非循环;无自圈;*v*和*v′*有唯一路径; *G-e*非连通; *G+e*唯一回路; *n*(*E*)=*n*-1)

是否存在？

生成森林 任意G都存在

生成树 连通图G

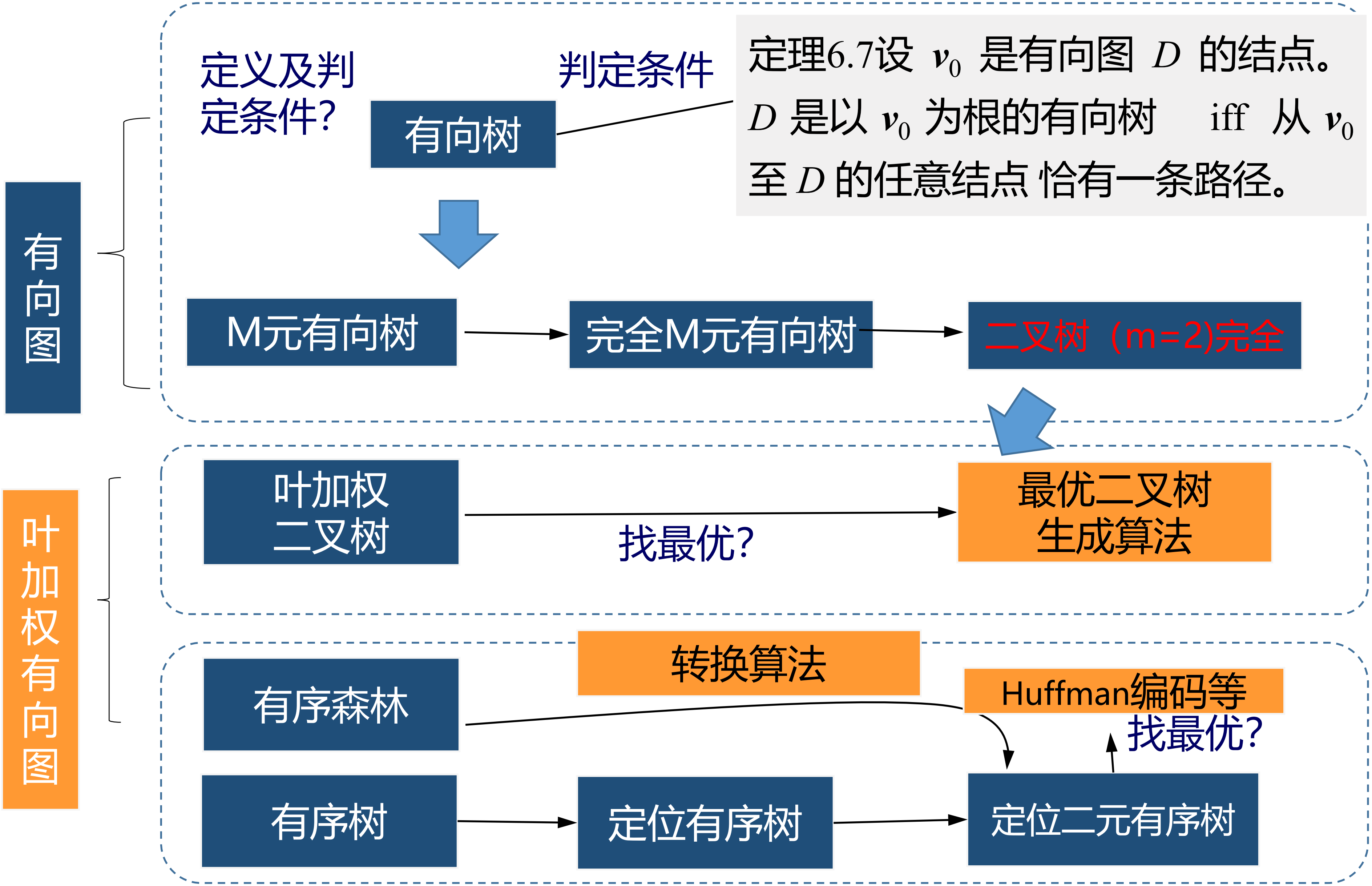
如何找？ N阶连通无向图 —避圈法 / 破圈法→ 树

加权图

找极值？ 最小生成树算法（MST）

D. 课堂思考

概念图谱

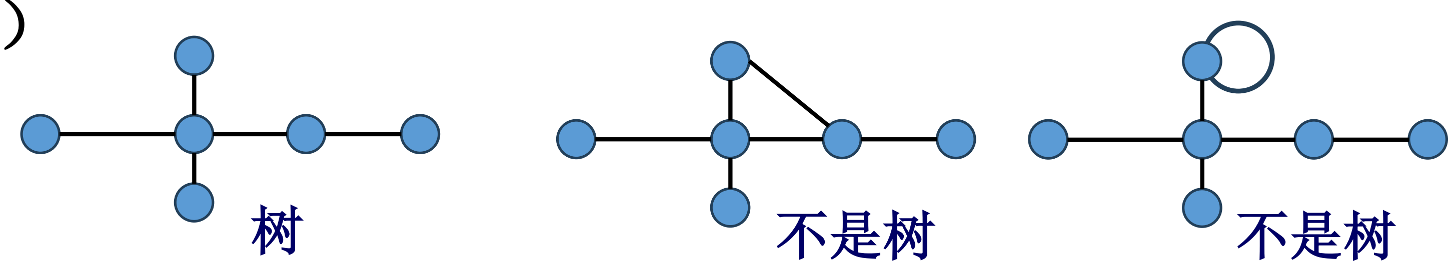

有向图
定义及判定条件?
有向树
判定条件
定理6.7设 v_0 是有向图 D 的结点。D 是以 v_0 为根的有向树 iff 从 v_0 至 D 的任意结点 恰有一条路径。
M元有向树
完全M元有向树
二叉树（m=2)完全
叶加权有向图
叶加权二叉树
找最优?
最优二叉树生成算法
转换算法
有序森林
Huffman编码等
找最优?
有序树
定位有序树
定位二元有序树

6.1 树

定义 6.1 非循环的连通无向图称为树。

- **非循环无向图**：没有回路的无向图
- **回路**：连通2度正则图

v到v的闭路径（除v外的其他结点和所有边恰好出现在闭路径上一次）



平凡树：只有一个顶点的无向图

叶子：树T中，度数为1的顶点称为叶子结点

分支顶点：树T中，度数大于1的顶点称为分支顶点

知识重点

树定义的等价条件

定理6.1 设 $G=\langle V, E, \Psi\rangle$ 是n阶无向图，则以下条件等价:

(1) G 是连通的和非循环的。

(2) G 无自圈，且当 $v, v' \in V$时，皆有唯一的一条从 v 至 v'的基本路径。

(3) G 是连通的，且当 $v, v' \in V$时，$e \notin E, \Psi' = \{\langle e, \{v, v'\}\rangle\}$时，$G+\{e\}_{\Psi'}$有唯一的一条回路。

(4) G 是连通的，且当$e \in E$时，$G-e$ 是非连通的。

(5) G 是连通的 且$n(E)=n-1$。

(6) G 是非循环的且有$n(E)=n-1$。

定理6.1 设 $G=\langle V, E, \Psi\rangle$ 是n阶无向图，则以下条件等价：

(1) G 是连通的和非循环的。

(2) G 无自圈，且当 $v, v'\in V$时，皆有唯一的一条从 v 至 v'的基本路径。

证明：(1)⇒ (2)

(a) 不存在自圈：由于 G是非循环的，因此G没有回路，所以 G 无自圈。

(b) 存在基础路径：由于 G 是连通的，因此 v可达 v'，从而存在一条从 v 至 v'的路径。

由定理3.1知，必存在一条从 v 至 v'的基本路径。

定理3.1 设图 $G=\langle V, E, \Psi\rangle$，$v, v'\in V$。如果存在从 v 至 v'的路径，则存在从 v 至 v'的基本路径。

证明： (1)⇒ (2)

(c) 基本路径的唯一性：假设基本路径不唯一，

P_1：$ve_1v_1e_2v_2 \ldots v_{p-1}e_pv'$，$P_2$：$ve'_1v'_1e'_2v'_2 \ldots, v'_{q-1}e'_qv'$，

是两条从 v 至 v'的不同的基本路径。

令G_1为以$\{v, v_1,\ldots, v_{p-1}, v'\}$为结点集，以$\{e_1, e_2,\ldots, e_p\}$为边集的子图（即$P_1$）；

G_2为以$\{v, v'_1,\ldots, v'_{q-1}, v'\}$为结点集，以$\{e'_1, e'_2,\ldots, e'_q\}$为边集的子图(即$P_2$)。

任取$e \notin E$，$\Psi' = \{< e, \{v, v'\} >\}$，

则$G_1+\{e\}_{\Psi'}$与$G_2+\{e\}_{\Psi'}$显然都是回路，从而是欧拉图，且可运算，也是 G 的子图。

证明： (1)⇒ (2)

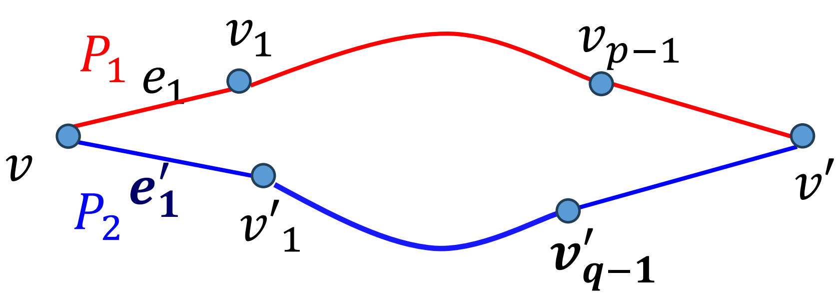
(c) 基本路径的唯一性：假设基本路径不唯一，

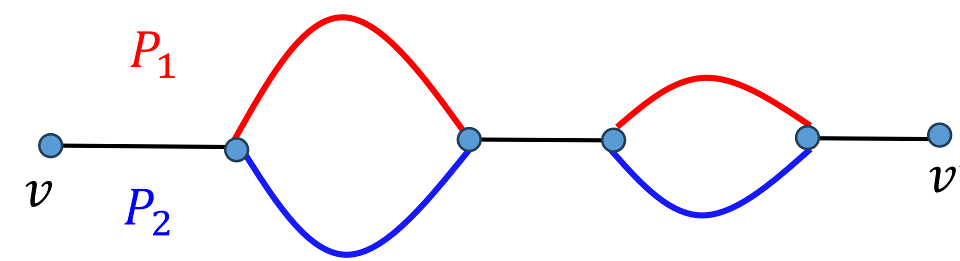
P_1：$ve_1v_1e_2v_2 \ldots v_{p-1}e_pv'$，$P_2$：$ve'_1v'_1e'_2v'_2 \ldots, v'_{q-1}e'_qv'$，

是两条从 v 至 v'的不同的基本路径。

令G_1为以{ $v, v_1,\ldots, v_{p-1}, v'$}为结点集，以{ $e_1, e_2,\ldots, e_p$}为边集的子图（即P_1）；

G_2为以{$v, v'_1,\ldots, v'_{q-1}, v'$}为结点集，以{$e'_1, e'_2,\ldots, e'_q$}为边集的子图(即$P_2$)。





证明： (1)⇒ (2)

(c) 基本路径的唯一性：假设基本路径不唯一，

P_1：$ve_1v_1e_2v_2 \ldots v_{p-1}e_pv'$，$P_2$：$ve'_1v'_1e'_2v'_2 \ldots, v'_{q-1}e'_qv'$，

是两条从 v 至 v'的不同的基本路径。

令G_1为以{ $v, v_1, \ldots, v_{p-1}, v'$}为结点集，以{ $e_1, e_2, \ldots, e_p$}为边集的子图（即P_1）；

G_2为以{$v, v'_1, \ldots, v'_{q-1}, v'$}为结点集，以{$e'_1, e'_2, \ldots, e'_q$}为边集的子图(即$P_2$)。

任取$e \notin E$，$\Psi' = \{< e, \{v, v'\} >\}$，

则$G_1 + \{e\}_{\Psi'}$与$G_2 + \{e\}_{\Psi'}$显然都是回路，从而是欧拉图，且可运算，也是 G 的子图。

证明： (1)⇒ (2)

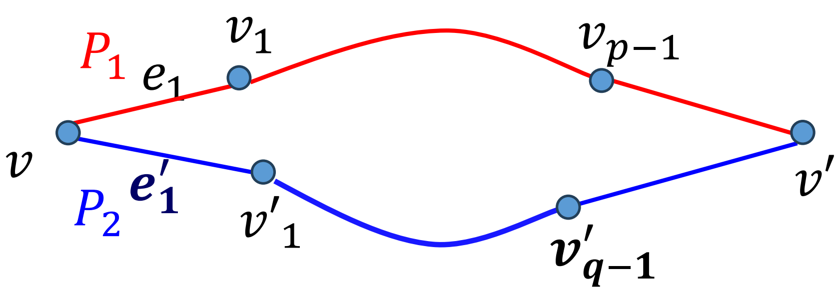
(c) 基本路径的唯一性（续）：

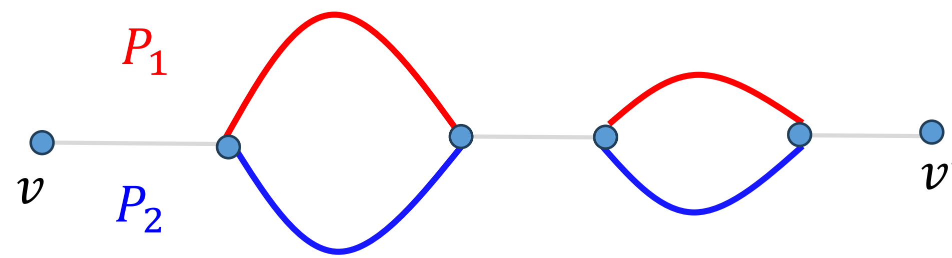
定理4.5 如果 G_1 和 G_2是可运算欧拉图，则$G_1 \oplus G_2$是欧拉图。

证明：(1)⇒ (2)

(c) <u>基本路径的唯一性（续）</u>：由定理4.5知，

$$G_1+\{e\}_{\Psi'}\oplus G_2+\{e\}_{\Psi'}$$ 是欧拉图。





由于P_1与P_2是两条不同的基本路径，

因此，$G_1+\{e\}_{\Psi'}\oplus G_2+\{e\}_{\Psi'}$不是零图。

所以，$G_1+\{e\}_{\Psi'}\oplus G_2+\{e\}_{\Psi'}$必有非平凡分支，记为 G''。

则G''中每个结点v的度 $d_{G''}(v)>1$。

由定理3. 9知，G''不是非循环图。

由于G''是G 的子图，得G 不是非循环图，与 G 是非循环图矛盾。

定理6.1 设 $G=\langle V, E, \Psi\rangle$ 是n阶无向图，则以下条件等价：

(2) G 无自圈，且当 $v, v' \in V$时，皆有唯一的一条从 v 至 v'的基本路径。

(3) G 是连通的，且当 $v, v' \in V$时，$e \notin E, \Psi'=\{\langle e, \{v, v'\}\rangle\}$时，$G+\{e\}_{\Psi'}$有唯一的一条回路。

证明：(2)⇒ (3)

(a) 连通性：由于对任意结点$v, v'\in V$，有从v 至 v'的基本路径，即v 可达v'，因此 G是连通的。

(b) 存在回路：当 $v, v' \in V$时，设从 v 至 v'的基本路径为：

$$P_1:\ v e_1 v_1 e_2 v_2 \dots v_{p-1} e_p v',$$

令G_1为以$\{v, v_1,\dots, v_{p-1}, v'\}$为结点集，以$\{e_1, e_2,\dots, e_p\}$为边集的子图（即$P_1$）。

取 $e \notin E, \Psi'=\{\langle e, \{v, v'\}\rangle\}$，$G_1+\{e\}_{\Psi'}$显然是回路，也是$G+\{e\}_{\Psi'}$的子图，因此，$G+\{e\}_{\Psi'}$有一条回路。

定理6.1 设 $G=<V, E, \Psi>$ 是n阶无向图，则以下条件等价：
(2) G 无自圈，且当 $v, v' \in V$时，皆有唯一的一条从 v 至 v'的基本路径。
(3) G 是连通的，且当 $v, v' \in V$时，$e \notin E, \Psi' = \{<e, \{v, v'\}>\}$时，$G+\{e\}_{\Psi'}$有唯一的一条回路。

证明： (2)⇒ (3)

(c) 回路的唯一性：假设C_1与C_2是$G+\{e\}_{\Psi'}$的两条不同的回路，则必是欧拉图，由定理4.5知，$C_1 \oplus C_2$ 为欧拉图。

由于C_1与C_1是不同的回路，因此$C_1 \oplus C_2$不是零图。

显然，$C_1 \oplus C_2$为 G 的子图，因此$C_1 \oplus C_2$ 必有非平凡分支，记为G'，其每个结点v的度 $d_{G'}(v)>1$。

由定理3.9知，G'不是非循环图，必有回路，记为 C。

因此，对C 中任意两个结点u和 v，必有两条不同的从u到v的基本路径，与(2)矛盾。

定理6.1 设 $G=<V, E, \Psi>$ 是n阶无向图，则以下条件等价：

(3) G 是连通的，且当 $v, v'\in V$时，$e\notin E$, $\Psi'=\{<e, \{v, v'\}>\}$时， $G+\{e\}_{\Psi'}$ 有唯一的一条回路。

(4) G 是连通的，且当$e\in E$时， $G-e$ 是非连通的。

证明： (3)⇒(4) 显然G是连通的。

下面用反证法证明当$e\in E$时， $G-e$ 是非连通的。

假设$G-e$ 是连通的。

设$\Psi(e)=\{u, v\}$，则由(3)知， $G=(G-e)+\{e\}_{\Psi}$ 有一条回路。

因此， G中必有两条从u到v的不同的基本路径。

由(1)⇒(2)的证明可知，G 必有回路C。

任取$e'\notin E$及$u\in V$，令$\Psi'=\{<e', \{u\}>\}$，即e'是一个自圈。

考虑 $G+\{e'\}_{\Psi'}$， 显然 C 也是它的一个回路。

因此， $G+\{e'\}_{\Psi'}$有两个不同的回路，与(3)矛盾。

定理6.1 设 $G=<V, E, \Psi>$ 是n阶无向图，则以下条件等价:

(4) G 是连通的，且当$e \in E$时， $G-e$ 是非连通的。

(5) G 是连通的 且$n(E)=n-1$。

证明： (4)$\Rightarrow$(5) 显然， G 是连通的。

可证，G 是连通的简单图。

否则，设G不是简单图，则

- G 有自圈e，或
- 存在两个结点$u, v \in V$，使得u到v有两条边e_1和e_2。

则 $G-e$, $G-e_1$, $G-e_2$均是连通的，与题设矛盾。

定理6.1 设 $G=<V, E, \Psi>$ 是n阶无向图，则以下条件等价：
(4) G 是连通的，且当$e \in E$时， $G-e$ 是非连通的。
(5) G 是连通的 且$n(E)=n-1$。

证明： (4)⇒ (5) 下面用关于n的第二数学归纳法证明$n(E)=n-1$。
(a) 当 $n=1$时， 由于 G 是简单图，因此没有边，即$n(E)=0$，结论成立。
(b) 假定对任意的$k \geq 2$， 当$n<k$ 时皆有$n(E)=n-1$;
设当 $n=k$ 时， 设 $n(E)=m$。
任取$e \in E$，由$G-e$是非连通图可知，$G-e$恰有两个分支G_1与G_2。
设$G_i\ (i=1,2)$ 有 n_i个节点和 m_i条边，根据归纳假设，必有

$$m_1=n_1-1, m_2=n_2-1。$$

从而得到 $m=m_1+m_2+1=n_1+n_2-1=k-1$,
因此，由归纳法证明了结论成立。

定理6.1 设 $G=<V,E,\Psi>$ 是n阶无向图，则以下条件等价:

(5) G 是连通的 且$n(E)=n-1$。

(6) G 是非循环的且有$n(E)=n-1$。

证明： (5)⇒(6) 用关于n的第一数学归纳法证明G是非循环图。

(a)当$n=1$时， $n(E)=0$，没有回路，故为非循环。

(b)假设$n=k$时，结论成立，即G是非循环的。

当$n=k+1$时，由于$n(E)=k$，由握手定理得，$\sum_{v\in V} d_G(v)=2n(E)=2k$。

由于 G为连通的，对每个$v\in V$，皆有$d_G(v)\geq 1$，

所以必存在$v'\in V$，使得$d_G(v')=1$；

否则$\sum_{v\in V} d_G(v)\geq 2k+2$，矛盾。

显然 $G-v'$ 是连通的；否则$G-v'$至少有两个分支，每个分支都有结点与v'关联，矛盾。

由于$G-v'$的阶为n-1=k且边数为$n(E)-1=k-1$，

根据归纳假设，$G-v'$必是非循环的。

由于$d_G(v')=1$，因此，G也必是非循环的。

定理6.1 设 $G = \langle V, E, \Psi \rangle$ 是n阶无向图，则以下条件等价:

(6) G 是非循环的且有$n(E) = n - 1$。

(1) G 是连通的和非循环的。

证明： (6)⇒ (1) 只需证明G连通的。

（反证法）假设 G 有 k 个分支$G_1, G_2, \ldots, G_k$ （$k>1$），

则每个$G_i(1 \leq i \leq k)$都是非循环且为连通的。

设$G_i(1 \leq i \leq k)$有 n_i个结点和 m_i 条边 。

树定义的等价条件

定理6.1 设 $G=\langle V, E, \Psi\rangle$ 是n阶无向图，则以下条件等价：

(1) G 是连通的和非循环的。

(2) G 无自圈，且当 $v, v'\in V$时，皆有唯一的一条从 v 至 v'的基本路径。

(3) G 是连通的，且当 $v, v' \in V$时，$e \notin E, \Psi'= \{\langle e, \{v, v'\}\rangle\}$ 时， $G+\{e\}_{\Psi'}$ 有唯一的一条回路。

(4) G 是连通的，且当$e \in E$时， $G-e$ 是非连通的。

(5) G 是连通的 且$n(E) = n-1$。

(6) G 是非循环的且有$n(E)= n-1$。

定理6.1 设 $G=<V, E, \Psi>$ 是n阶无向图，则以下条件等价：

(6) G 是非循环的且有$n(E)= n-1$。

(1) G 是连通的和非循环的。

证明： (6)⇒ (1) 只需证明G连通的。

（反证法）假设 G 有 k 个分支$G_1, G_2,\dots, G_k$ （$k>1$），

则每个$G_i(1 \le i \le k)$都是非循环且为连通的。

设$G_i(1 \le i \le k)$有 n_i个结点和 m_i 条边 。

所以由前面的论证知道必有 (1)⟹ (5)，

因此$m_i = n_i - 1(1 \le i \le k)$，

从而得到$n - 1 = \sum_{i=1}^{k} m_i = \sum_{i=1}^{k}(n_i - 1) = n - k$，

得 $k = 1$，即 G 只有一个分支，

因此，G必是连通的。

树的等价定义

定理6.1 设 $G=<V, E, \Psi>$ 是n阶无向图，则以下条件等价：

(1) G 是连通的和非循环的。

(2) G 无自圈，且当 $v, v' \in V$时，皆有唯一的一条从 v 至 v'的基本路径。

(3) G 是连通的，且当 $v, v' \in V$时，$e \notin E, \Psi' = \{<e, \{v, v'\}>\}$时，$G+\{e\}_{\Psi'}$有唯一的一条回路。

(4) G 是连通的，且当$e \in E$时，$G-e$ 是非连通的。

(5) G 是连通的 且$n(E)=n-1$。

(6) G 是非循环的且有$n(E)=n-1$。

定理 6.2 阶大于 1 的树至少有两个端点。

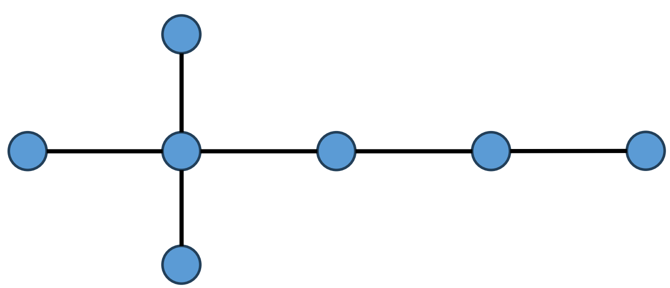
- 度为 1 的结点 称为 端点 （**定义**1. 7）

叶子：树 T中，度数为1的顶点称为叶子结点

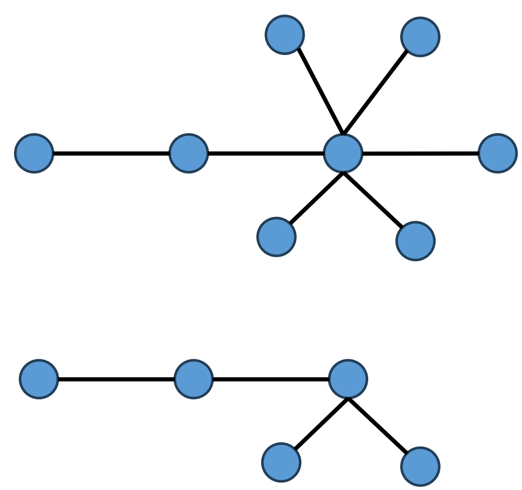
6.2 森林和生成树

■ 树是非循环的连通无向图，如果去掉对连通性的要求，就得到森林的概念。

定义6.2 每个分支都是树的无向图称为森林。



树即是森林



森林

定理 6.3 如果森林 F 有 n 个结点，m 条边和 k 个分支，则 $m=n-k$。

$n-m$ 棵树

树

证明：由定理6.1知，n 个顶点的树有 n-1 条边。

设每个分支有 n_i 个顶点，则：

$$n_1+n_2+\ldots+n_k = n。$$

因此，森林一共有

$$(n_1\text{-}1)+(n_2\text{-}1)+\ldots+(n_k\text{-}1) = n\text{-}k = m$$

条边。

所以 $m=n\text{-}k$。

生成树(Spanning Tree)、生成森林

定义 6.3

(1) 如果树 T是无向图 G的生成子图，则称 T为 G的生成树。

(2) 如果森林 F 是无向图 G的生成子图，则称 F 为 G的生成森林。

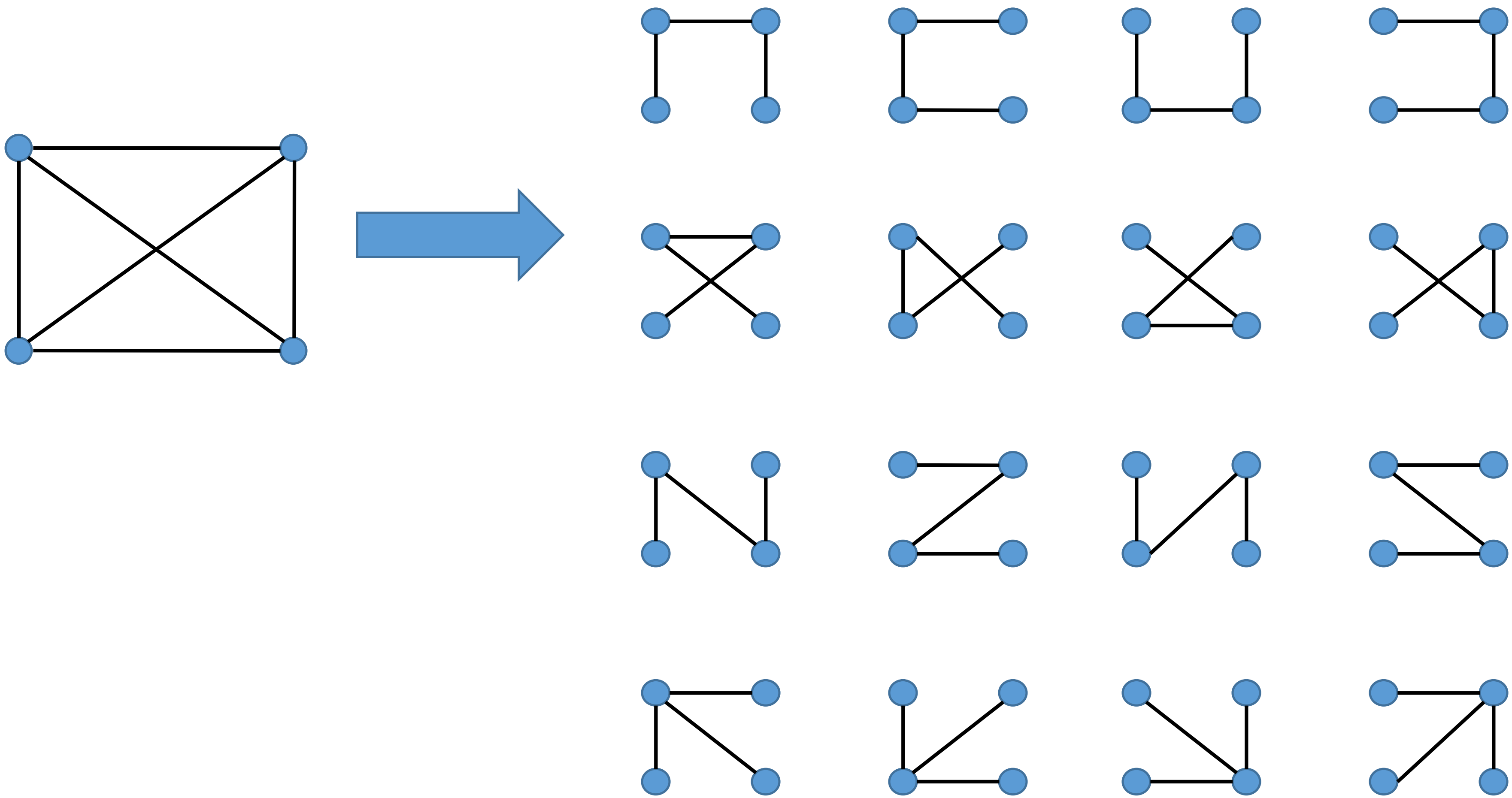
- 显然有以下结论：

定理 6.4

(1) 每个无向图都有生成森林。

(2) 无向图G 有生成树当且仅当G是连通的。

生成树(Spanning Tree)示例



连通图的生成树构造方法

- **避圈法：**

  添加 e_1，…，e_i，在添加的每一步均保证：e_{i+1}不与 $\{e_1, \ldots, e_i\}$ 的任何子集构成回路。

- **破圈法：**

  在 G_0 (即 G) 中去掉 e_1 得到 G_1，

  在 G_1 中去掉 e_2 得到 G_2，

  在 G_2 中去掉 e_3 得到 G_3，…

其中 e_i 为 G_{i-1} 中某条回路中的边，直到没有回路，即把 G 中的所有回路均挑破

定理：设无向图G连通，则G至少有一个生成树。

该定理的证明过程实际上是求生成树的算法：

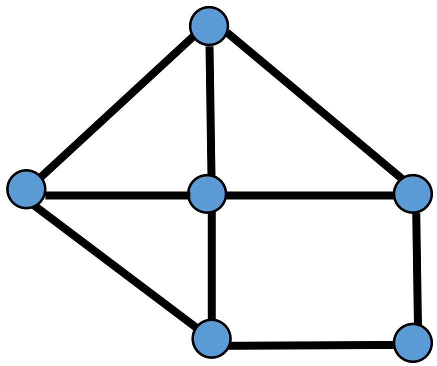
输入：连通无向图G

输出：生成树 T_G

1. $i \leftarrow 0$，$G_0 \leftarrow G$;
2. 若G_i无圈，则$T_G \leftarrow G_i$并终止；否则转（3）；
3. 找出G_i中任何一圈α_i，并从α_i中去掉任何一边e_i，
   $G_{i+1} \leftarrow G_i - e_i$;
4. $i \leftarrow i+1$，转（2）

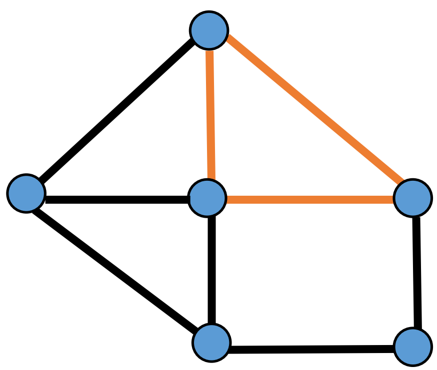
“破圈法”：逐次破掉图G中所有的圈，并保证每破一圈时都得到G的一个连通生成子图，因而最后得到的T_G保证是G的生成树。

用破圈法求图G的生成树

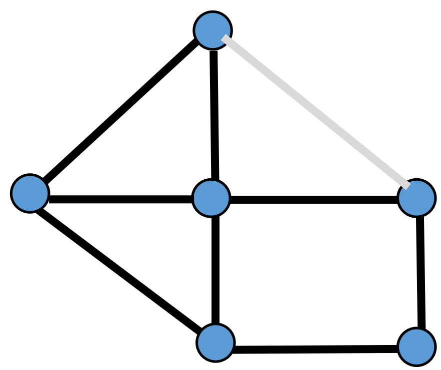


G_0

用破圈法求图G的生成树

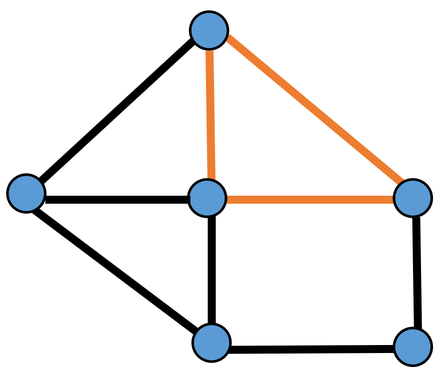


G_0

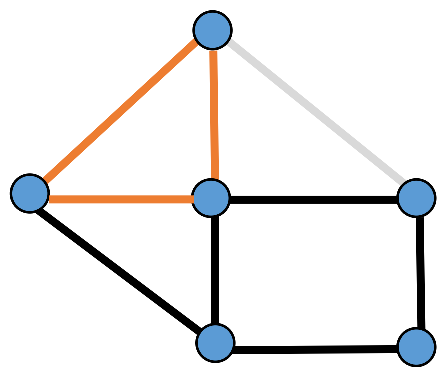


G_1

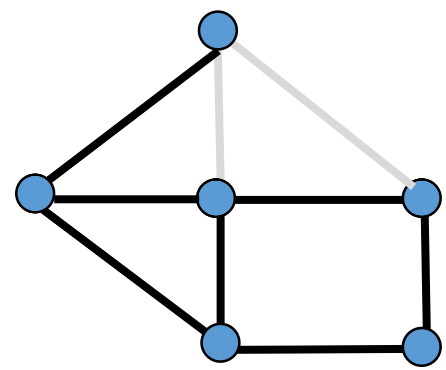
用破圈法求图G的生成树



G_0

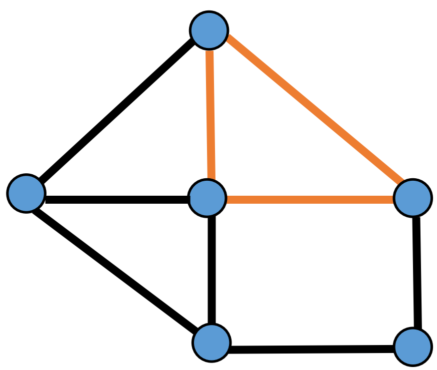


G_1

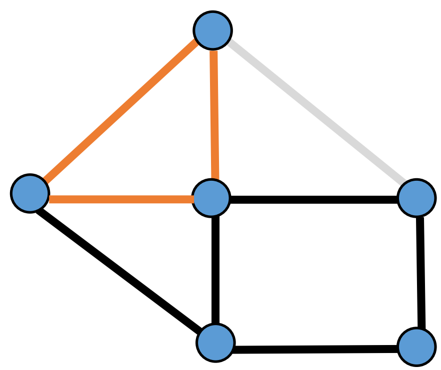


G_2

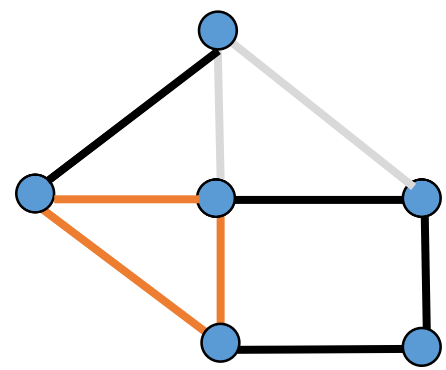
用破圈法求图G的生成树



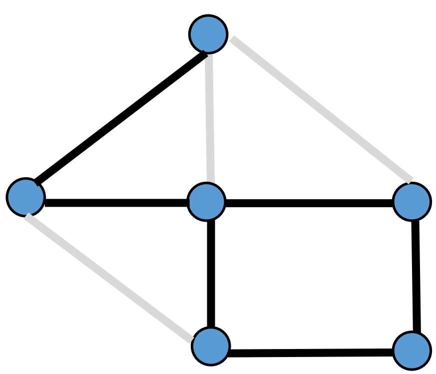
G_0



G_1

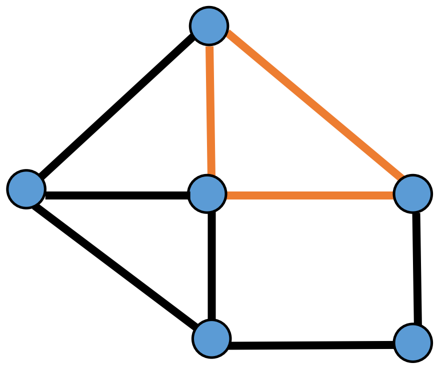


G_2

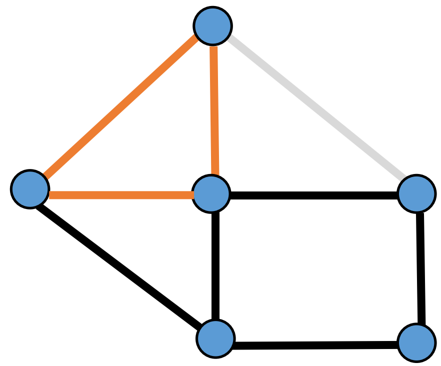


G_3

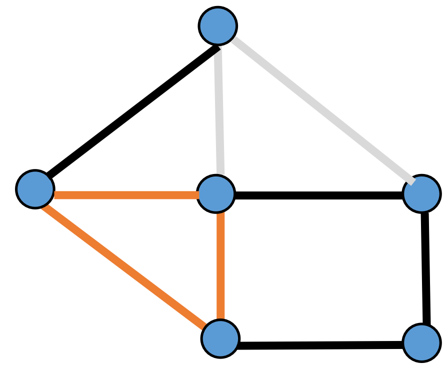
用破圈法求图G的生成树



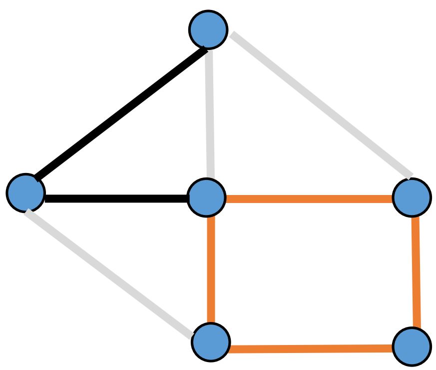
G_0



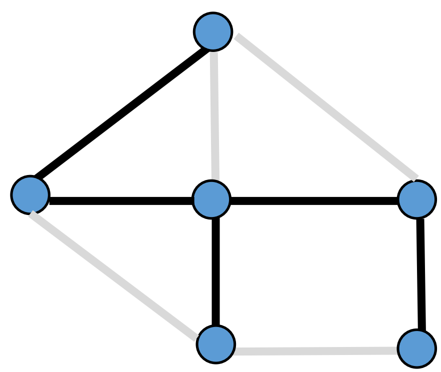
G_1



G_2



G_3



G_4

最小生成树-Minimum Spanning Tree (MST)

定义6. 4 (1)设 $<G, W>$是加权图，$G' \subseteq G$。G'中所有边的加权长度之和 称为 G'的加权长度。

(2) 设 G 是连通无向图，$<G, W>$是加权图，G的所有生成树中加权长度最小者称为$<G, W>$的最小生成树。

贪心法求解最小生成树常用的有两种算法:

(1) Prim's MST algorithm (prim算法).

(2) Kruskal's MST algorithm(kruskal算法).

Prim算法是基于点的，而Kruskal算法是基于边的。

最小生成树求法（避圈法、破圈法）

按 “避圈法” 求最小生成树：

设 G 是有 m 条边的 n 阶连通无向图，

1. 把 G 的 m 条边按加权长度递增的顺序排成 $e_1, e_2, \ldots, e_m$;
2. $T \leftarrow \varnothing$;
3. $j \leftarrow 1$，$i \leftarrow 1$;

（i记录正在扫描的边的下标；j记录T中边数是否已达n-1）

4. 若 $j = n$ 则算法结束。
5. 若 G 的以 $T \cup \{e_i\}$ 为边集合的子图没有回路，

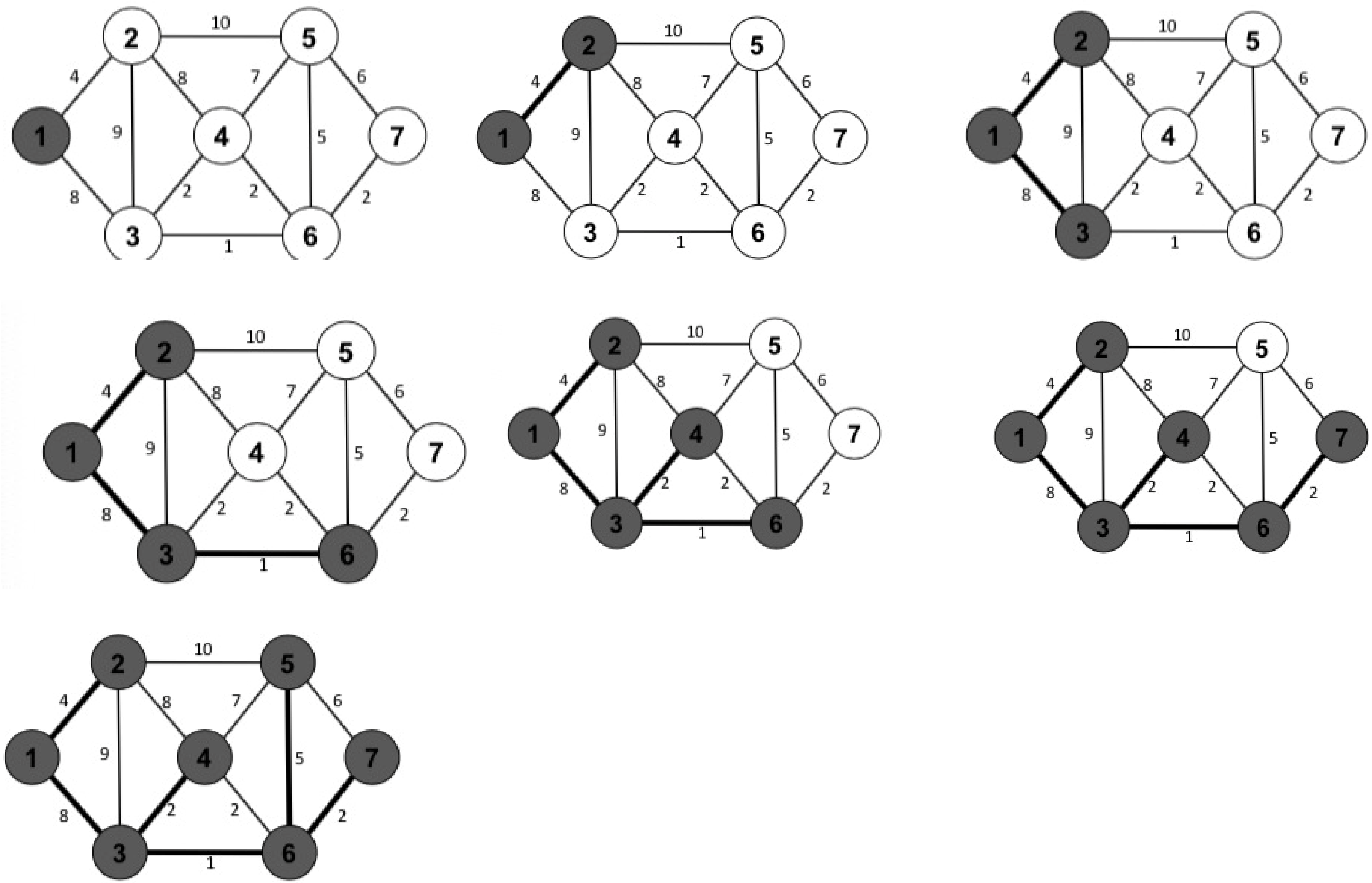
   则 $T \leftarrow T \cup \{e_i\}$ 且 $j \leftarrow j+1$;
6. $i \leftarrow i+1$，转向 4;

算法结束时，T 即为所求的最小生成树的边集。

最小生树算法1--Prim算法

- 用于连通无向图，贪心算法
  1. 维护Tree结构
  2. 初始$E=\{\}$, $V=\{v\}$ //任取节点v
  3. 循环n-1次
     - 选择一条边(v_1, v_2), 满足
       - $v_1 \in V$, $v_2 \notin V$,
       - (v_1, v_2)权值最小
     - $E=E \cup (v_1, v_2)$
     - $V=V \cup \{v_2\}$

Prim算法：示例

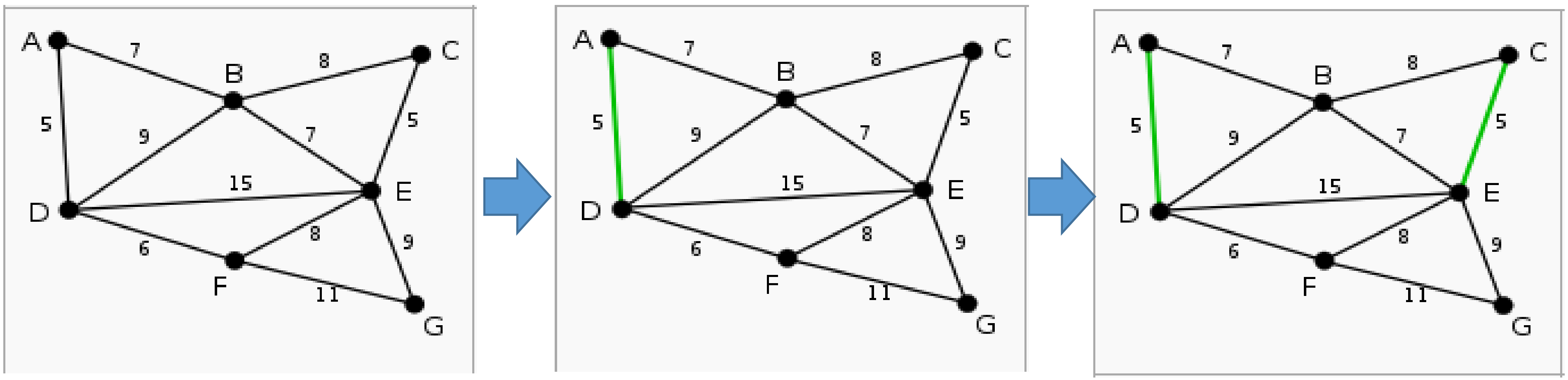


最小生树算法2--Kruskal算法

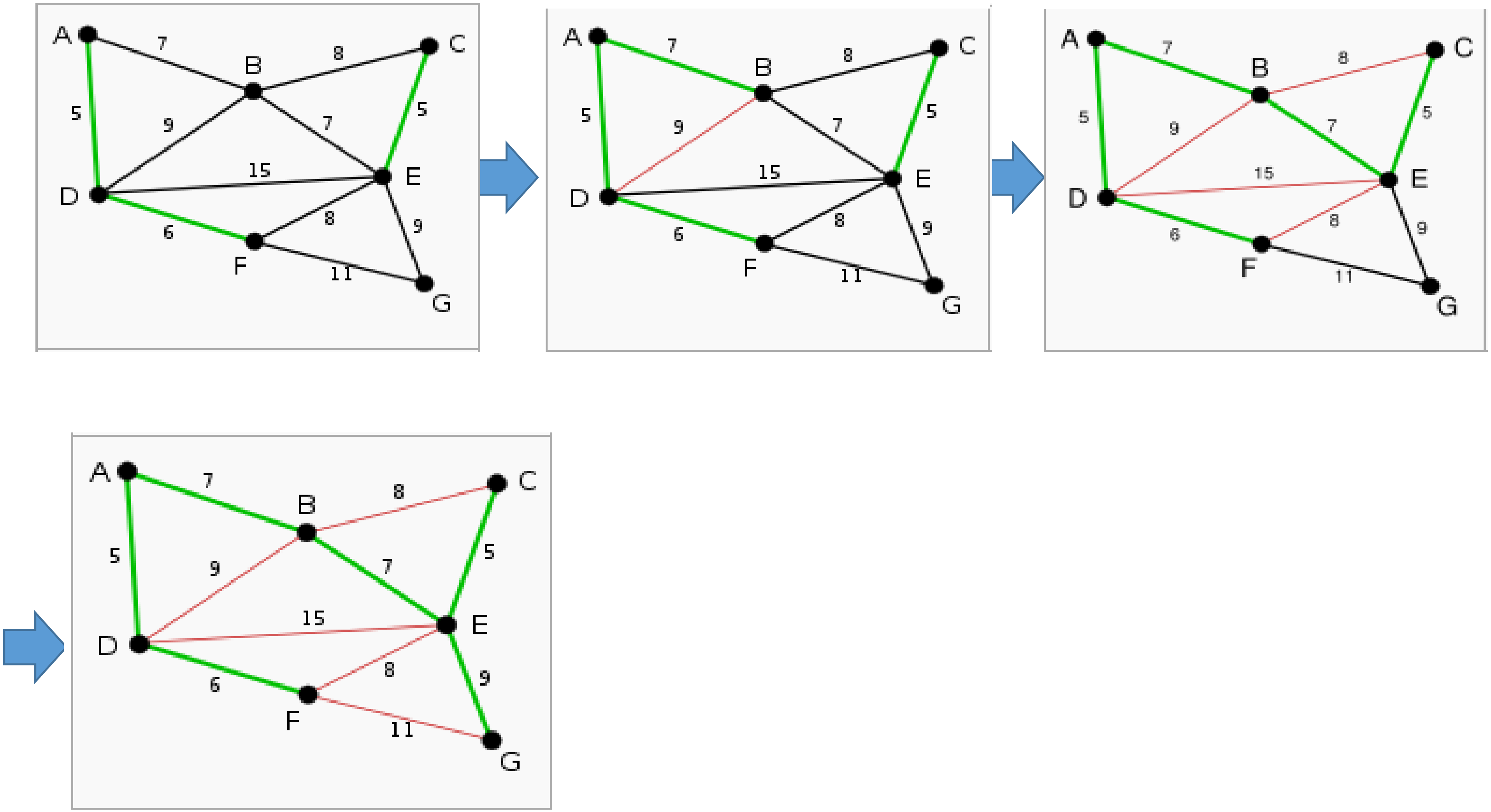
- 贪心算法
  - 将边按权重从小到大排序
  - 按顺序选择每条边，只要与已选择边不构成圈，就选择。
  - 终止条件
    - ✓已经选择了n-1条边；
    - ✓如果处理所有边，仍然不够n-1条，则说明图不连通

D. 课堂思考 如何判断圈？并查集，*quick union*

Kruskal算法举例



Kruskal算法举例



两种算法对比

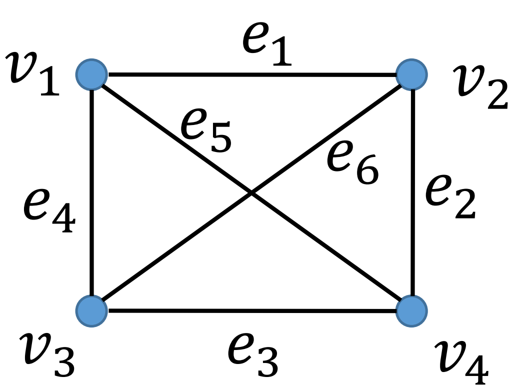
		Prim算法	对比 Kruskal算法
概念	操作	点操作	边操作，更简单
	查找	直接查找 多次寻找邻边权重最小值	贪心策略， 需要先对权重排序后查找
应用	计算复杂性	适合稠密图, O(n2)	适合稀疏图，O(elog(e))
	效率	需对邻边多次排序	一般更快，对所有边一次排序

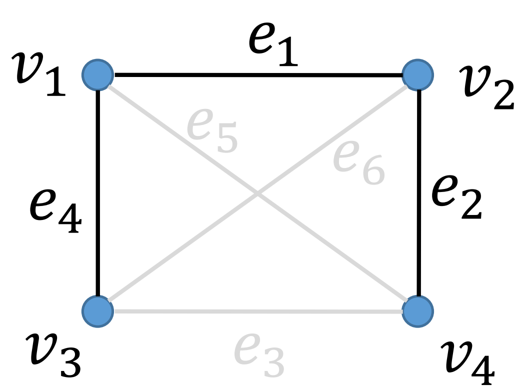
连通无向图：枝、弦

定义 6.5 设 T 是<u>连通无向图</u> G 的生成树，称 T 的边为枝，而 G 的不属于 T 的边称为 弦。

问题：连通图 G 的边 e 是枝还是弦？

- 与给定的生成树 T 密切相关。

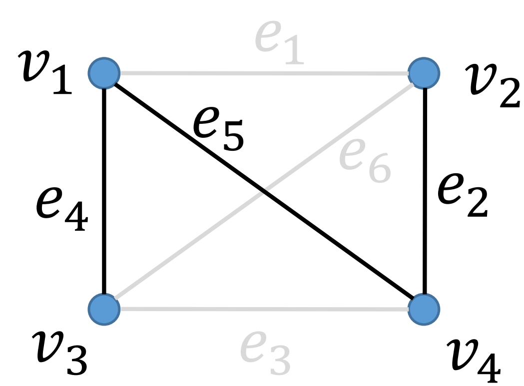




T_1

枝： e_1, e_2, e_4

弦： e_3, e_5, e_6



T_2

枝： e_2, e_4, e_5

弦： e_1, e_3, e_6

连通无向图：枝、弦

定义 6.5 设 T 是连通无向图 G 的生成树，称 T 的边为枝，而 G 的不属于 T 的边称为弦。

问题：连通图 G 的边 e 是枝还是弦？

- 与给定的生成树 T 密切相关。
- 对于G的某个生成树 T， e 是枝，而对于 G 的另一个生成树 T_1， e 却可能是弦。
- 但是，对于 G 的任何生成树，枝的数目和弦的数目都是固定的。

定理6.5 设 G 是有 m 条边的 n 阶连通无向图，则对于 G 的任何生成树 T，都有 $n-1$ 个枝和 $m-n+1$ 个弦。

无向图：圈秩、余圈秩

定理6.5 设 G 是有 m 条边的 n 阶连通无向图，则对于 G 的任何生成树 T，都有 $n-1$ 个枝和 $m-n+1$ 个弦。

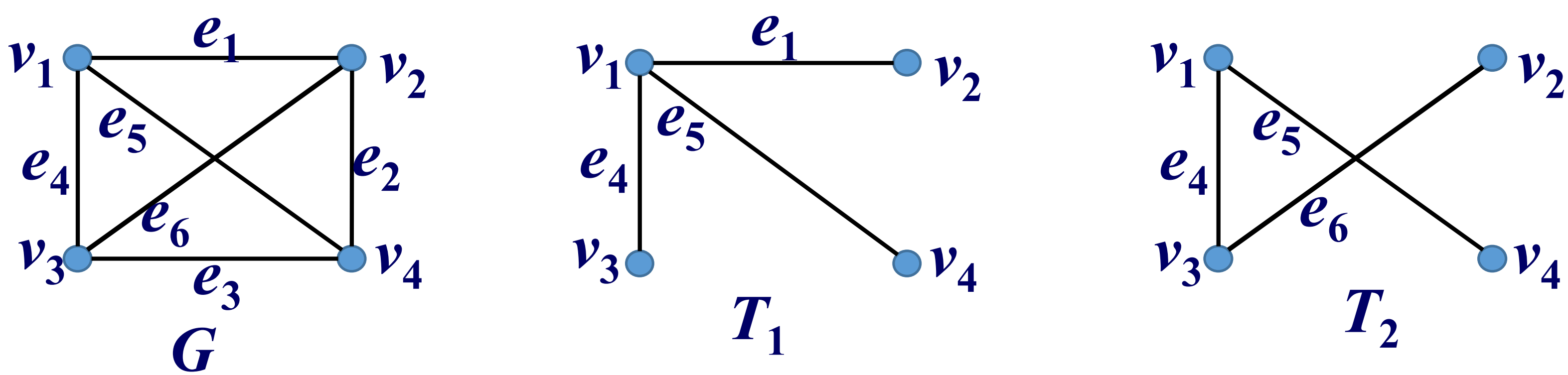
定义 6.6 若 n 阶无向图 G 有 m 条边和 k 个分支，则 G 的余圈秩 $r = n - k$，圈秩 $\mu = m - n + k$。

- 如果 G 是连通图（$k = 1$），则 G 的余圈秩 r 是枝的数目， 圈秩 μ 是弦的数目。

基本回路（圈）

- 由定理6.1知，如果在生成树中增加一条弦，则恰产生一个回路。

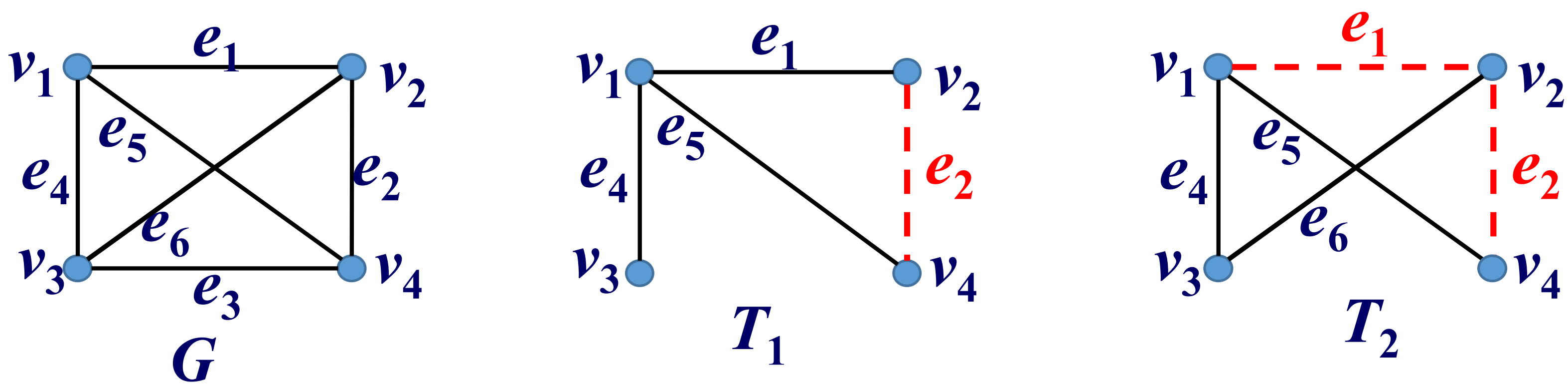
定义6.7（基本回路）设 T 是连通无向图 G 的生成树，G 的只包含一条弦的回路称为基本回路。



基本回路（圈）

- 由定理6.1知，如果在生成树中增加一条弦，则恰产生一个回路。

定义6.7（基本回路）设 T 是连通无向图 G 的生成树，G 的只包含一条弦的回路称为基本回路。



- 基本回路的概念与生成树相关联
  - 某回路对这个生成树是基本回路，而对另一个生成树却未必是基本回路。
  - 对于给定连通图的任何生成树，基本回路的数目都是相同的

定理 6.6 设 T 是连通无向图 G 的任意生成树。

(1) 基本回路的数目等于 G 的圈秩 μ； $\mu = m-(n-1)$

(2) 对于G的任意回路 C，总可以找到若干个基本回路 $C_1, C_2, \ldots, C_k$，使 C 与 $C_1 \oplus C_2 \oplus \ldots \oplus C_k$ 的差别仅在于孤立点。

证明：(1) 显然。（一条弦，对应一个基本回路）

(2) 设 C 是 G 的任意回路且C包含 k 条弦，显然 k>0，

定义 6.6 若 n 阶无向图 G 有 m 条边和 k 个分支，则 G 的余圈秩 $r = n - k$，圈秩 $\mu = m - n+k$。

定理 6.6 设 T 是连通无向图 G 的任意生成树。

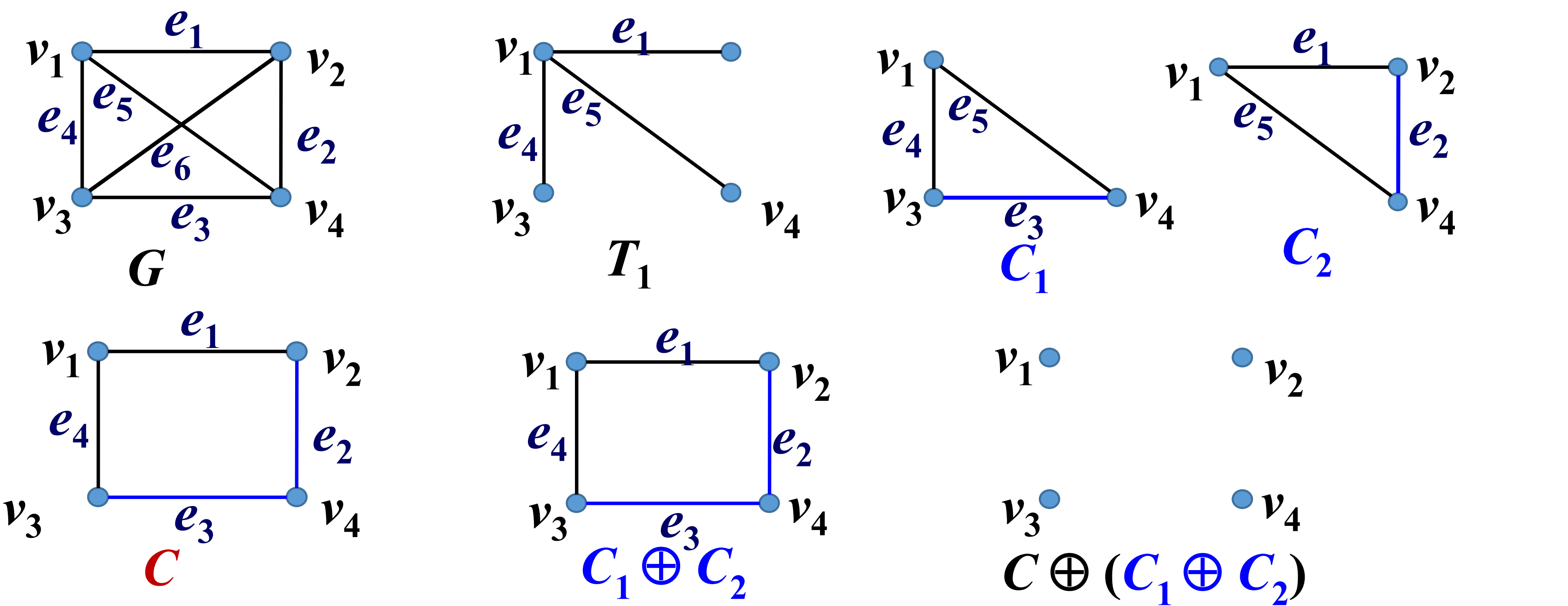
(1) 基本回路的数目等于 G 的圈秩 μ；

$\mu = m-(n-1)$

(2) 对于G的任意回路 C，总可以找到若干个基本回路 $C_1, C_2, \ldots, C_k$，使 C 与 $C_1 \oplus C_2 \oplus \ldots \oplus C_k$ 的差别仅在于孤立点。

证明：(1) 显然。（一条弦，对应一个基本回路）

(2) 设 C 是 G 的任意回路且C 包含 k 条弦，显然 $k>0$，



定理 6.6 设 T 是连通无向图 G 的任意生成树。

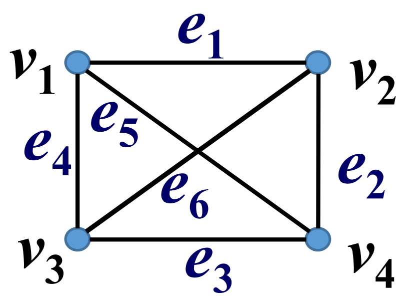
(1) 基本回路的数目等于 G 的圈秩 μ；

$\mu = m-(n-1)$

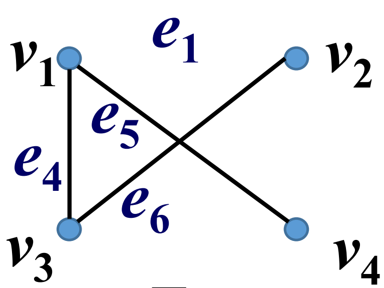
(2) 对于G的任意回路 C，总可以找到若干个基本回路 $C_1, C_2, \ldots, C_k$，使 C 与 $C_1 \oplus C_2 \oplus \ldots \oplus C_k$ 的差别仅在于孤立点。

证明：(1) 显然。（一条弦，对应一个基本回路）

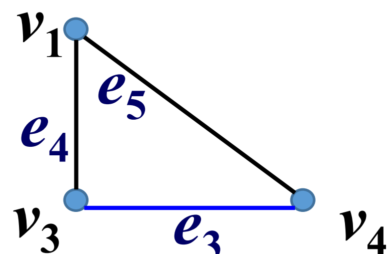
(2) 设 C 是 G 的任意回路且C 包含 k 条弦，显然 k>0，



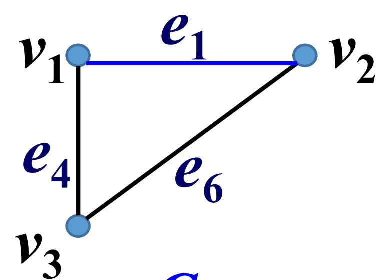
G



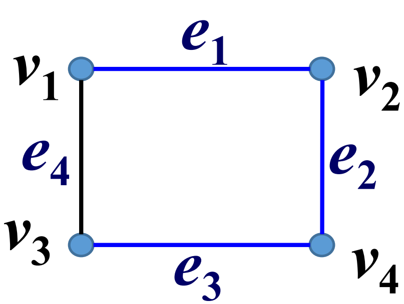
T_1



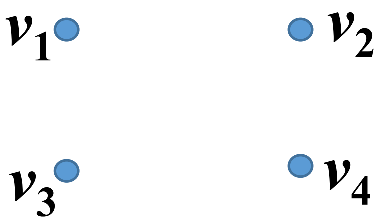
C_1



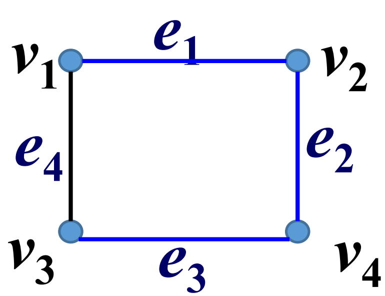
C_2



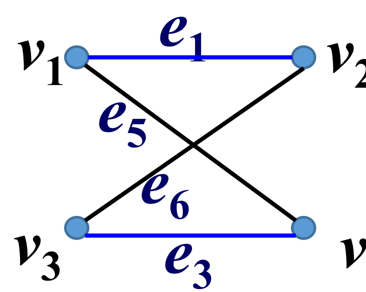
C



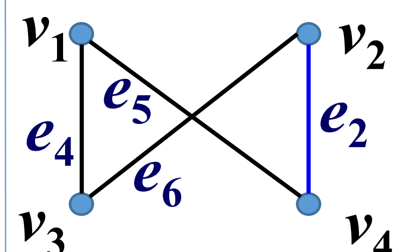
$C\oplus(C_1\oplus C_2\oplus C_3)$



$C_1\oplus C_2\oplus C_3$



$C_1 \oplus C_2$



C_3

定理 6.6 设 T 是连通无向图 G 的任意生成树。

(1) 基本回路的数目等于 G 的圈秩 μ； $\mu = m-(n-1)$

(2) 对于G的任意回路 C，总可以找到若干个基本回路 $C_1, C_2, \ldots, C_k$，使 C 与 $C_1 \oplus C_2 \oplus \ldots \oplus C_k$ 的差别仅在于孤立点。

证明：(1) 显然。（一条弦，对应一个基本回路）

(2) 设 C 是 G 的任意回路且C包含 k 条弦，显然 $k>0$，

设这 k 条弦是 $e_1, e_2, \ldots, e_k$ ，

C_i 是包含 e_i 的基本回路 (i =1，2，…，k)。

令 $C' = C_1 \oplus C_2 \oplus \ldots \oplus C_k$,

则 C' 包含的弦 也是 e_1，e_2，…，e_k。

因此，$C \oplus C'$ 中的边都是枝，则 $C \oplus C'$ 是非循环的 。

定义 6.6 若 n 阶无向图 G 有 m 条边和 k 个分支，

则 G 的余圈秩 $r = n - k$，圈秩 $\mu = m - n + k$。

定理 6.6 设 T 是连通无向图 G 的任意生成树。
(1) 基本回路的数目等于 G 的圈秩 μ；
(2) 对于G的任意回路 C，总可以找到若干个基本回路 $C_1, C_2, \ldots, C_k$，使 C 与 $C_1 \oplus C_2 \oplus \ldots \oplus C_k$ 的差别仅在于孤立点。

证明（续）下面证明 $C \oplus C'$ 是零图。
若 $C \oplus C'$ 不是零图，必有一分支是阶大于 1 的树，
根据定理 6.2， $C \oplus C'$ 有端点。
另，因为 C 和 C' 都是欧拉图，所以$C \oplus C'$ 是欧拉图。
这与 $C \oplus C'$ 有端点矛盾，故 $C \oplus C'$ 必为零图，
即C 与 C'的差别仅在于孤立点。

6.3 有向树

定义 6.8 一个结点的入度为 0，其余结点的入度均为 1 的弱连通有向图称为有向树。其中，

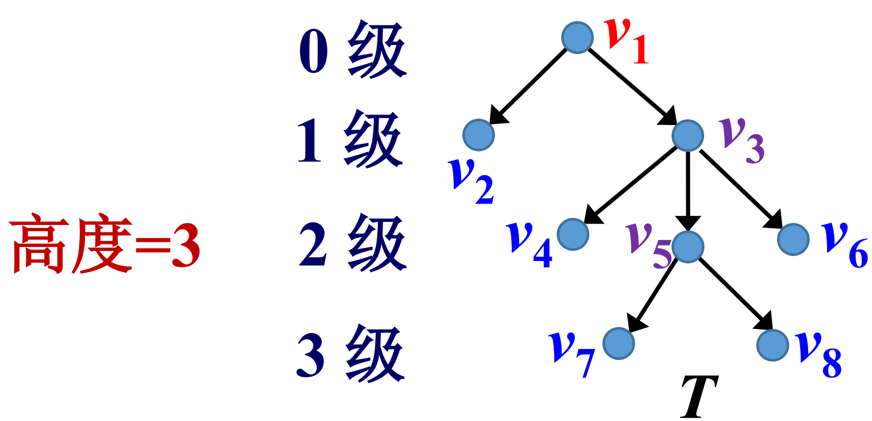
(1) 入度为 0 的结点称为根，

(2) 出度为 0 的结点称为叶，

(3) 出度大于 0 的结点称为分支结点，

(4) 从根至任意结点的距离称为该结点的级，

(5) 所有结点的级的最大值称为有向树的高度。



根：v_1

叶：v_2，v_4，v_6, v_7 , v_8

分支结点：v_1，v_3，v_5

定理 6.7 设 v_0 是有向图 D 的结点。 D 是以 v_0 为根的有向树当且仅当从 v_0 至 D 的任意结点恰有一条路径。

证明: (必要性) 设 $D = \langle V, E, \Psi \rangle$ 是有向树，v_0 是D的根。

因为 D是弱连通的，任取 $v' \in V$，则存在从 v_0 至 v'的半路径 P，

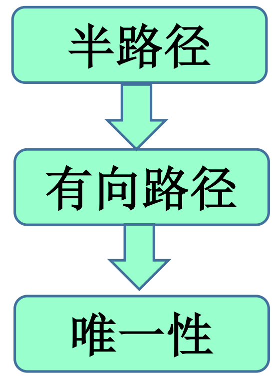
设 P为 $v_0 e_1 v_1 \ldots v_{p-1} e_p v_p$，其中 $v_p = v'$。

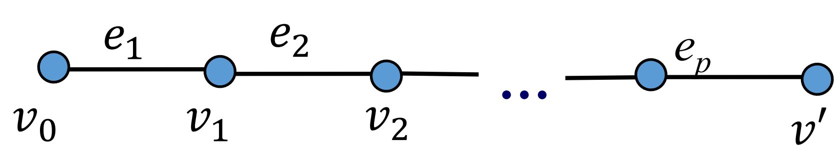
因为 $d_D^-(v_0) = 0$，所以 e_1 是正向边；

因为$d_D^-(v_1) =1$，所以 e_2 也是 正向边。

由归纳法可以证明： 每个 e_i（$1 \le i \le p$）均是正向边。

故 P 为有向路径。





定理 6.7 设 v_0 是有向图 D 的结点。 D 是以 v_0 为根的有向树当且仅当从 v_0 至 D 的任意结点恰有一条路径。

证明: (必要性) 设 $D = \langle V, E, \Psi \rangle$ 是有向树，v_0 是D的根。

因为 D是弱连通的，任取 $v' \in V$，则存在从 v_0 至 v'的半路径 P，

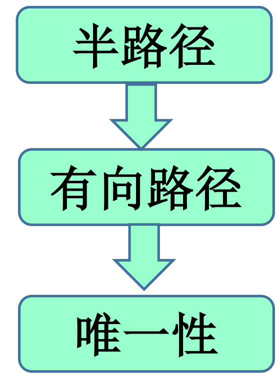
设 P为 $v_0 e_1 v_1 \ldots v_{p\text{-}1} e_p v_p$，其中 $v_p = v'$。

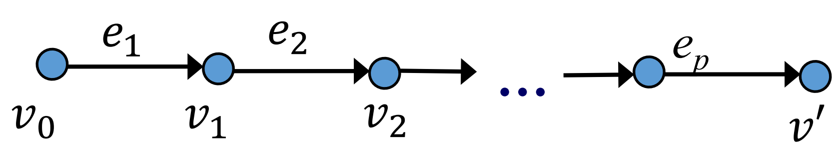
因为 $d_D^-(v_0) = 0$，所以 e_1 是正向边；

因为$d_D^-(v_1) =1$，所以 e_2 也是 正向边。

由归纳法可以证明： 每个 e_i（$1 \leq i \leq p$）均是正向边。

故 P 为有向路径。





定理 6.7 设 v_0 是有向图 D 的结点。D 是以 v_0 为根的有向树当且仅当从 v_0 至 D 的任意结点恰有一条路径。

证明: (必要性) 设 $D=\langle V, E, \Psi\rangle$ 是有向树，v_0 是D的根。

因为 D是弱连通的，任取 $v'\in V$，则存在从 v_0 至 v'的半路径 P，

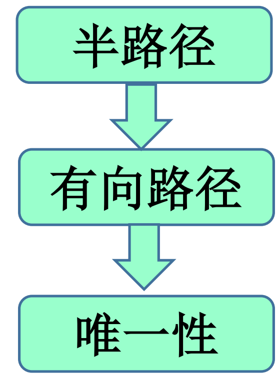
设 P为 $v_0 e_1 v_1 \ldots v_{p\text{-}1} e_p v_p$，其中 $v_p=v'$。

因为 $d_D^-(v_0)=0$，所以 e_1 是正向边；

因为$d_D^-(v_1)=1$，所以 e_2 也是 正向边。

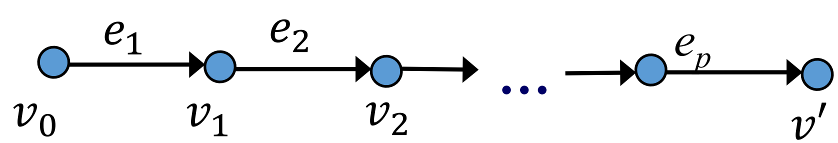
由归纳法可以证明：每个 e_i（$1\le i\le p$）均是正向边。

故 P 为有向路径。

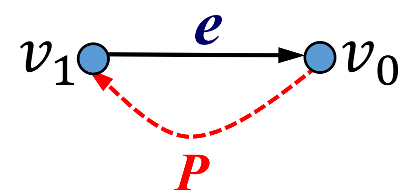


若从v_0至v'有两条路径 P_1 和 P_2，则P_1和P_2至少有一个公共点的入度大于1，与 D 是有向树矛盾。

故从从v_0至v'恰有一条路径。



定理 6.7 设 v_0 是有向图 D 的结点。D 是以 v_0 为根的有向树当且仅当从 v_0 至 D 的任意结点恰有一条路径。



证明: (充分性) 显然，D 是弱连通的。

(1) 若$d^-_D(v_0)>0$，则存在边 e 以v_0为终点。

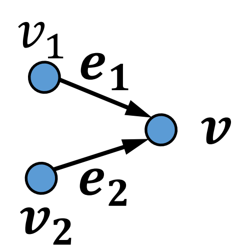
设v_1是e的起点，P是从v_0至v_1的路径，

则在D中存在<u>两条不同的</u>从v_0至v_0的路径：Pv_1ev_0和$Pv_1ev_0Pv_1ev_0$，

与已知条件矛盾，所以 $d^-_D(v_0)=0$。

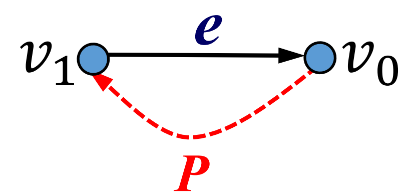
(2) 若 $d^-_D(v)>1$，其中v是D的节点，则存在两条边e_1和e_2以v为终点。

设e_1和e_2的起点分别是v_1和v_2，



定理 6.7 设 v_0 是有向图 D 的结点。D 是以 v_0 为根的有向树当且仅当从 v_0 至 D 的任意结点恰有一条路径。

证明: (充分性) 显然，D 是弱连通的。



(1) 若 $d_D^-(v_0) > 0$，则存在边 e 以 v_0 为终点。

设 v_1 是 e 的起点，P 是从 v_0 至 v_1 的路径，

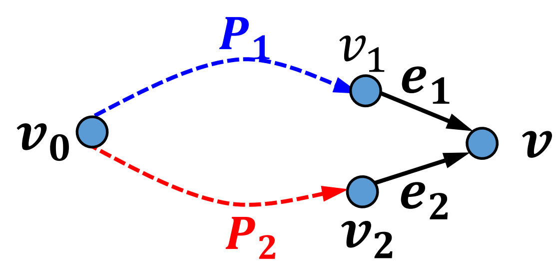
则在 D 中存在<u>两条不同的</u>从 v_0 至 v_0 的路径：Pv_1ev_0 和 $Pv_1ev_0Pv_1ev_0$，

与已知条件矛盾，所以 $d_D^-(v_0) = 0$。

(2) 若 $d_D^-(v) > 1$，其中 v 是 D 的节点，则存在两条边 e_1 和 e_2 以 v 为终点。

设 e_1 和 e_2 的起点分别是 v_1 和 v_2，从 v_0 至 v_1 和从 v_0 至 v_2 的路径分别是 P_1 和 P_2，则 P_1e_1v 和 P_2e_2v 是两条不同的从 v_0 至 v 的路径，与已知条件矛盾。

所以，D 是有向树，且 v_0 是 D 的根。



有向树的归纳定义

定义6.9 有向树归纳定义如下：

(1) 平凡图是有向树，其结点称为该有向树的根。

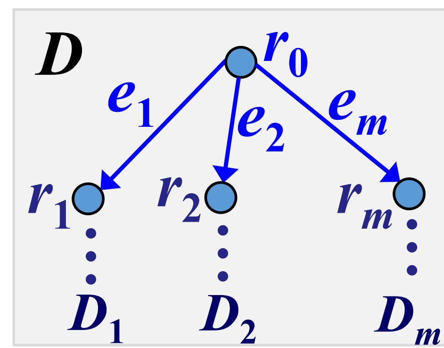
(2) 设 $m \in I_+$, $D_1, D_2, \ldots, D_m$ 分别是以 $r_1, r_2, \ldots, r_m$ 为根的有向树，并且两两不相交，

无公共结点、无公共边

r_0 不是 $\cup_{i=1}^{m} D_i$ 的结点，$e_1, e_2, \ldots, e_m$ 不是 $\cup_{i=1}^{m} D_i$ 中的边，且

$$\Psi: \{e_1, e_2, \ldots, e_m\} \to \{r_0, r_1, \ldots, r_m\}^2$$

定义为 $\Psi(e_i) = \langle r_0, r_i \rangle$ $(i=1, 2, \ldots, m)$。



若 $G = \langle \{r_0, r_1, \ldots, r_m\}, \{e_1, e_2, \ldots, e_m\}, \Psi \rangle$，则

$D = G \cup (\cup_{i=1}^{m} D_i)$ 是有向树，r_0 是 D 的根，并且称 $D_1, D_2, \ldots, D_m$ 是 D 的子树。

定义2.4 设 $G=<V, E, \Psi>$， $G'=<V', E', \Psi'>$同为无向图或同为有向图。

(1)如果对于任意 $e \in E \cap E'$，均有 $\Psi(e) = \Psi'(e)$，

则称 G 和 G'是可运算的。

(2)如果$V \cap V' = E \cap E' = \varnothing$，则称 G 和 G'是不相交的。

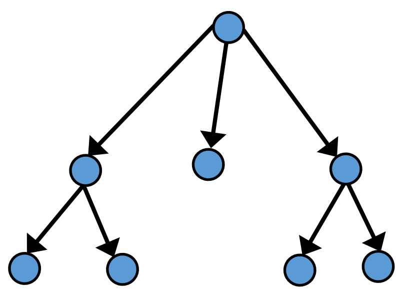
- 无公共结点、无公共边

(3)如果 $E \cap E' = \varnothing$，则称 G 和 G'是边不相交的。

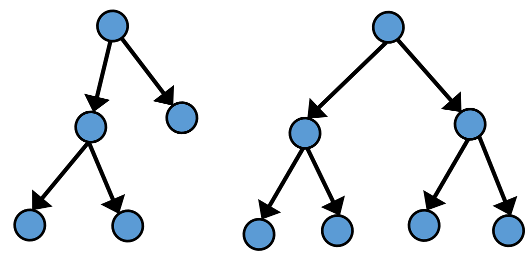
- 无公共边，可能有公共结点

有向森林

定义6.10 每个弱分支都是有向树的有向图，称为有向森林。



有向树即是有向森林



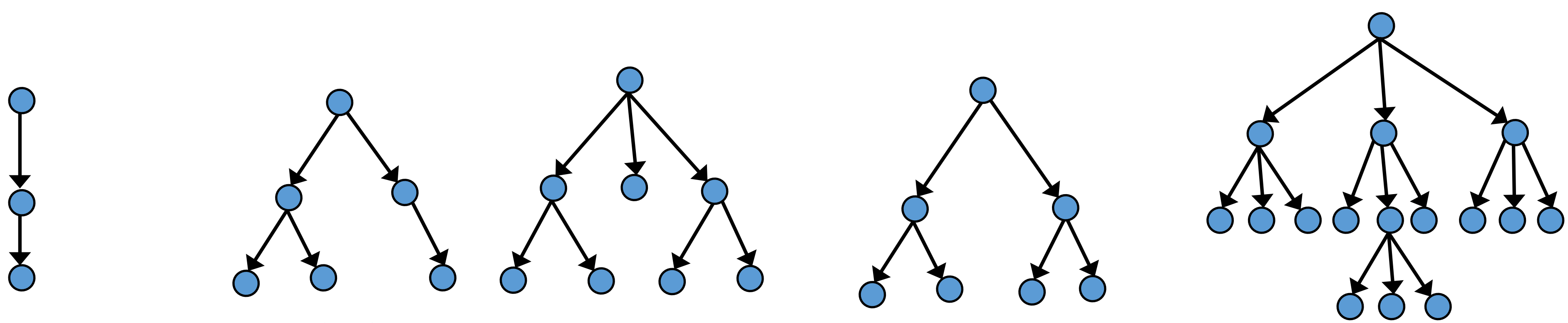
有向森林

(完全)m元有向森林

定义6. 11 设 $m\in\mathbf{N}$，D 为有向树。

(1)如果D的所有结点出度的最大值为m，则称D为m元有向树。

(2)如果对于m元有向树 D 的每个结点v, 皆有$d_D^+(v)=m$ 或 $d_D^+(v)=0$，则称D为完全m元有向树。



一元有向树　二元有向树　三元有向树　完全二元有向树　完全三元有向树

完全一元有向树

(完全)m元有向森林

定义6.11 设 $m \in \mathbf{N}$，D 为有向树。

(1)如果D的所有结点出度的最大值为m，则称D为m元有向树。

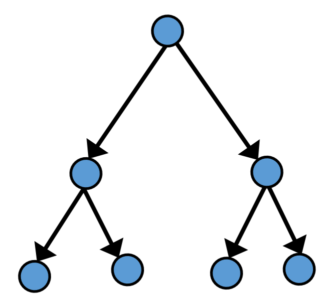
(2)如果对于m元有向树 D 的每个结点v, 皆有$d_D^+(v) = m$ 或 $d_D^+(v) = 0$，则称D为完全m元有向树。

- 完全二元有向树也称二叉树。

用途：字母和符号识别程序 {+， –， *， / }

00 01 10 11

统计字母出现的频繁程度



完全二元有向树

两个问题

- 编码问题
  - 假设 A B C D 四个字母：如何编码？
  - 出现频率是0.5, 0.3, 0.05, 0.15，如何编码？

知识点1：叶加权二叉树，Huffman编码

- 树的存储计算
  - 二叉树具有特点和良好性质；
  - 对于不同类型树，是否能统一存储和计算？

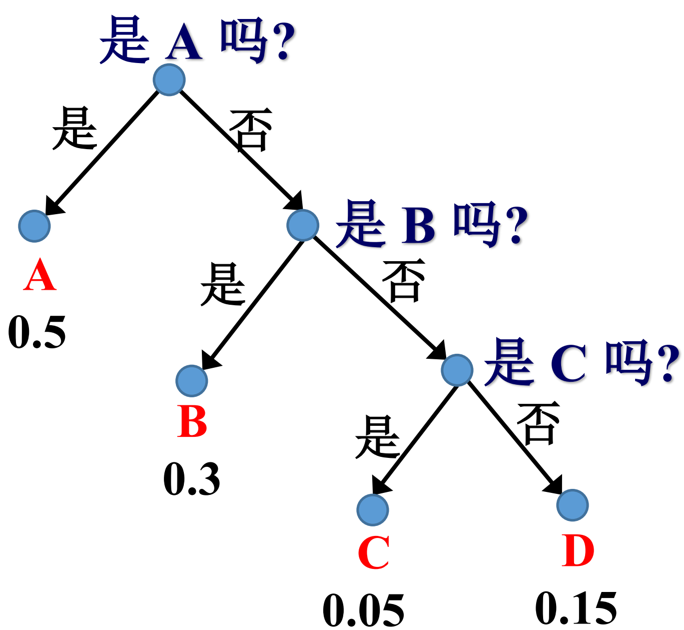
知识点2：森林-树-二叉树的转化

叶加权二叉树->最优二叉树->求解

定义6.12 (1) 设 V 是二叉树 D 的叶的集合， W: $V \to R_+$，则称 $<D, W>$为叶加权二叉树。

(2) 对于 D 的任意叶 v，称 $W(v)$ 为 v 的权，

称 $\sum_{v\in V}(W(v)\cdot L(v))$ 称为$<D, W>$的叶加权路径长度，其中 $L(v)$为 v的级。



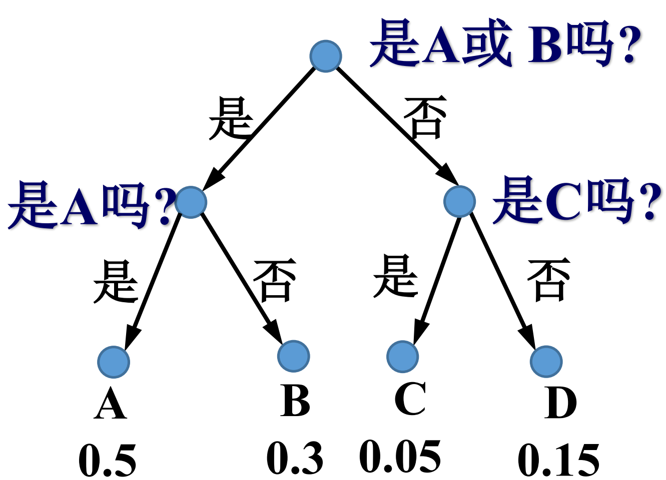
- 用叶表示字母或符号，
- 用分支结点表示判断，
- 用权表示字母或符号出现的概率，
- 则叶加权路径长度就表示算法的平均执行时间。

叶加权路径长度为：

$0.5\cdot 1+0.3\cdot 2+0.05\cdot 3+0.15\cdot 3=1.7$

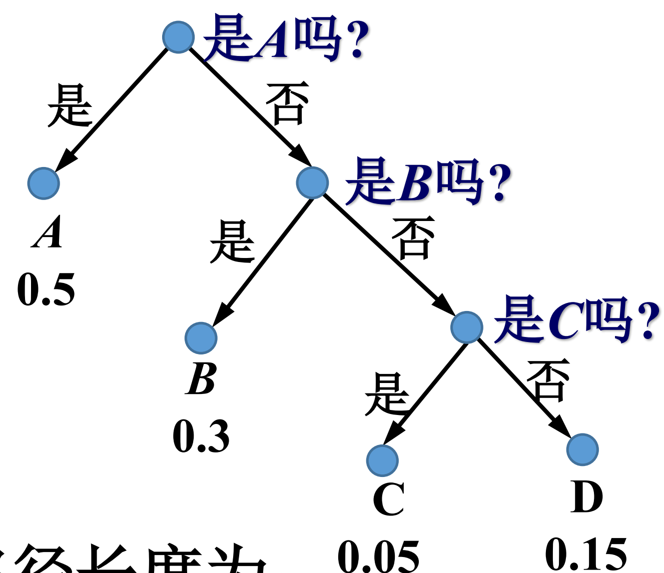
叶加权二叉树->最优二叉树->求解

定义 6.13 设 $<D, W>$ 是叶加权二叉树。如果对任一叶加权二叉树 $<D', W'>$，只要对于任意正实数 r，D 和 D' 中权等于 r 的叶的数目相同，就有 $<D, W>$ 的叶加权路径长度不大于 $<D', W'>$ 的叶加权路径长度，则称 $<D, W>$ 为最优的。



叶加权路径长度为:

$0.5 \cdot 2+0.3 \cdot 2+0.05 \cdot 2+0.15 \cdot 2 =2$



叶加权路径长度为:

$0.5 \cdot 1+0.3 \cdot 2+0.05 \cdot 3+0.15 \cdot 3 =1.7$

最优二叉树的递归求解过程

- 假设找有m个叶，其权分别是$w_1, w_2, \ldots, w_m$的最优二叉树。
  不妨设$w_1 \le w_2 \le \ldots \le w_m$。

设$<D, W>$是满足要求的最优二叉树，D中以$w_1, w_2, \ldots, w_m$为权的叶分别为$v_1, v_2, \ldots, v_m$。

权越小的叶的级越大

显然，在所有叶中，v_1, v_2的级最大，

不妨设v_1, v_2与同一个分支结点v'邻接。

令$D' = D-\{v_1, v_2\}\cup\{v'\}$，$W': \{v', v_3, v_4, \ldots, v_m\}\to R_+$，满足

$$W'(v') = w_1 + w_2, \quad W'(v_i) = w_i, \quad i=3, 4, \ldots, m,$$

易证：$<D, W>$是最优的当且仅当$<D', W'>$是最优的。

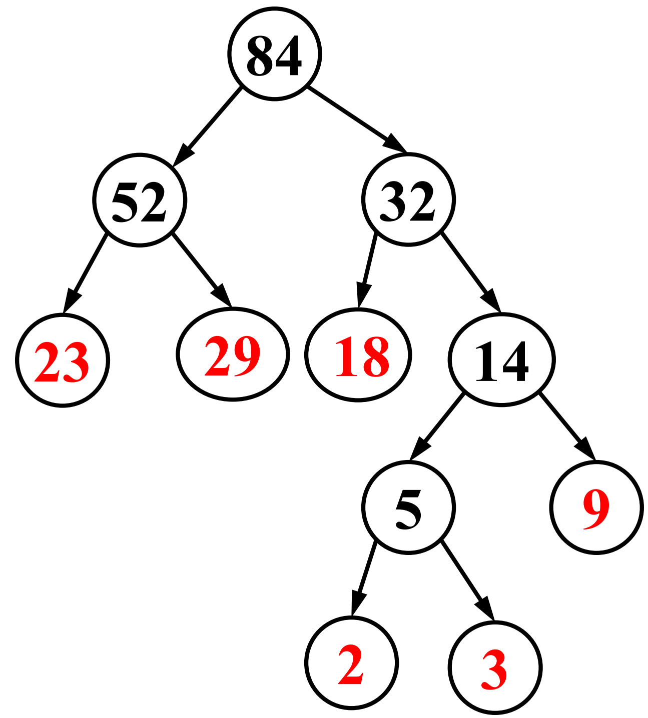
把求<u>m个叶</u>的最优二叉树归结为求<u>m-1个叶</u>的最优二叉树。

继续以上过程，直至归结为求2个叶的最优二叉树。

叶加权二叉树->最优二叉树->求解

- 最优二叉树求取算法：（举例说明）

2	3	9	18	23	29
	5	9	18	23	29
		14	18	23	29
			32	23	29
			32		52
					84



叶加权路径长度：

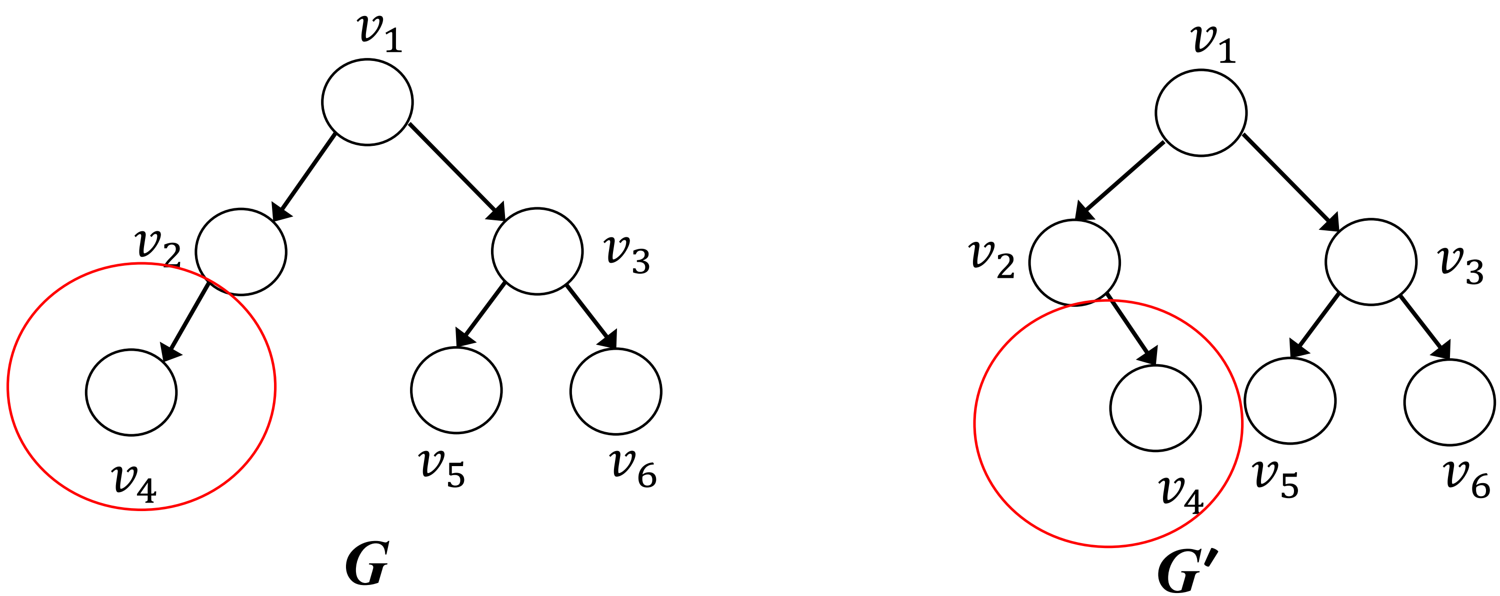
$$2\cdot 4+3\cdot 4+9\cdot 3+18\cdot 2+23\cdot 2+29\cdot 2 = 187$$

- 所有分支结点的数值之和就是叶加权路径长度

出度>0的结点

叶加权路径长度：84+52+32+14+5 = 187

两棵树是否是同一棵树



- G与G'是相同的有序树，因为同一级上结点的次序相同。
- 如果考虑结点之间的相对位置，G与G'不相同。
- G与G'是不同的定位有序树。

有序树/有序森林->定位有序树

定义 6.14

(1) 为每一级上的结点规定了次序的有向树称为有序树。

(2) 如果有向森林 F 的每个弱分支都是有序树，并且也为 F 的每个弱分支规定了次序，则称 F 为有序森林。

- 在画有序树时，总是把根画在上部，并规定同一级上结点的次序是从左至右。
- 在画有序森林时，弱分支的次序也是从左至右。

有序树/有序森林->定位有序树

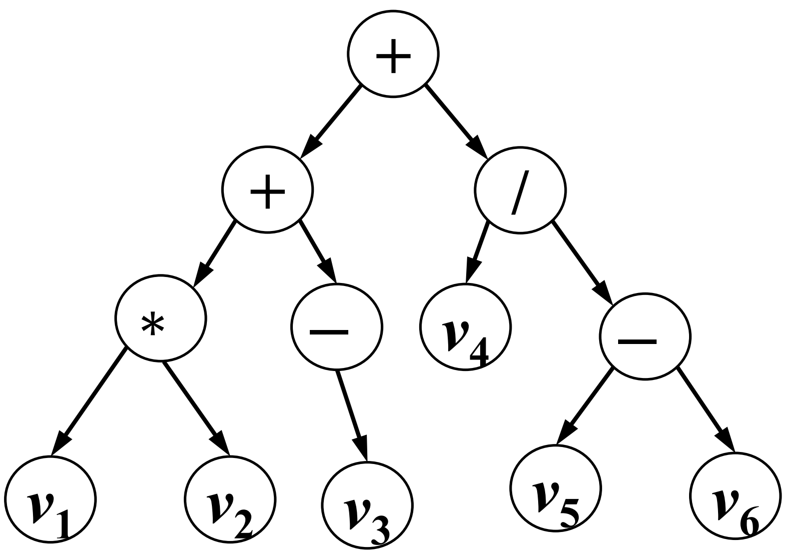
例：可以用有序树表示算术表达式，其中叶表示参加运算的数或变量，分支结点表示运算符。

$((v_1 * v_2) + (-v_3)) + v_4 / (v_5 - v_6)$

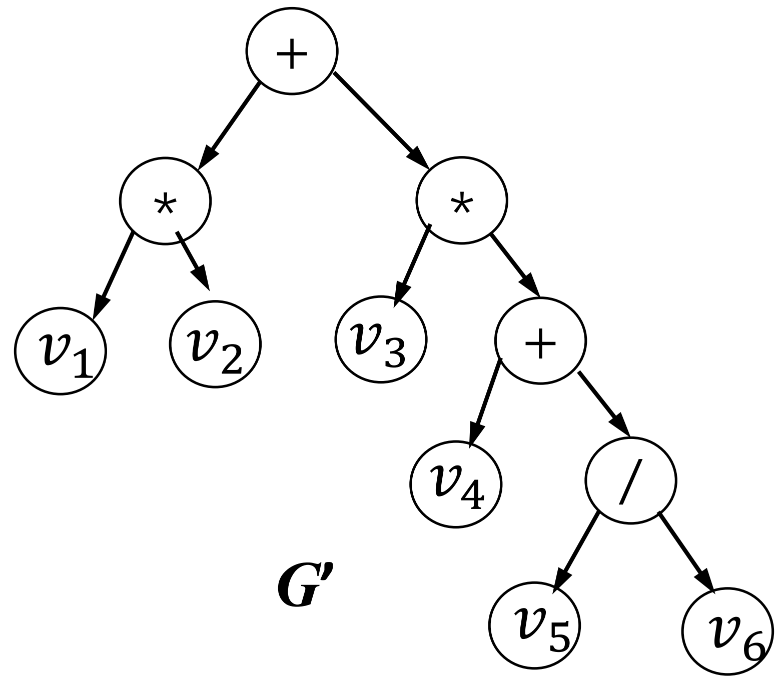
可表示为图G的有序数。

$v_1 * v_2 + v_3 * (v_4 + v_5 / v_6)$

可表示为图G'的有序树。

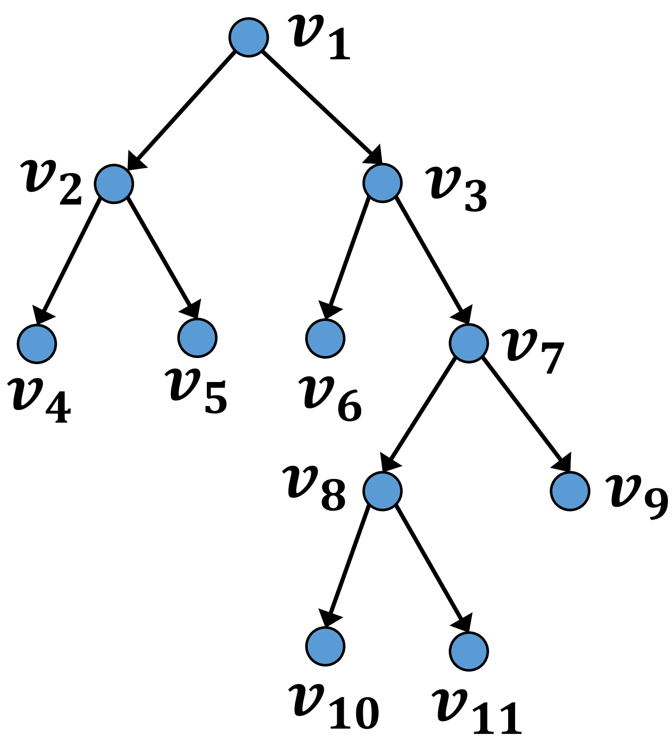


G



G'

有序树的结点



v_1是v_2与v_3的父亲

v_2与v_3是v_1的儿子

v_1是v_4与v_5的祖先

v_4与v_5是v_1的后代

v_2是v_1的左儿子

v_3是v_1的右儿子

有序树/有序森林->定位有序树

定义6.15 为每个分支结点的儿子规定了位置的有序树称为定位有序树。

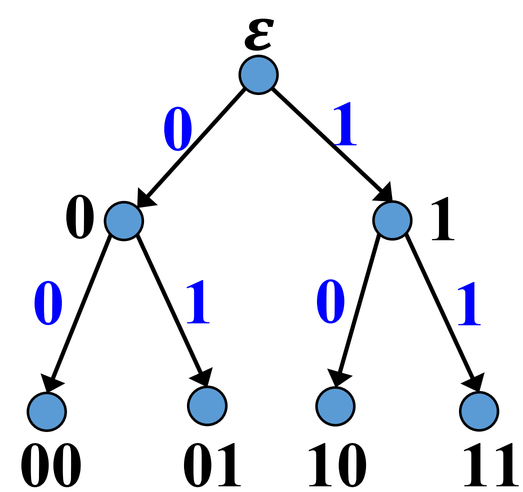
出度>0的结点

例：在定位二元有序树中，可用字符表{0, 1}上的字符唯一地表示每个结点表示二进制编码情况。

(1)用空子府串ε表示根；

(2)设用 β 表示某分支结点，则用β_0 表示它的左儿子，用β_1表示它的右儿子。

这样，每个结点都有了唯一的编码表示，并且不同结点的编码表示不同。



定位二元有序树全体叶的编码表示集合称为它的前缀编码

例：在计算机通信中要传输A, B, C, D, E, F, G, H 八个字母，它们出现频率为 A: 30%, B: 20%, C: 15%, D: 10%, E: 10%, F: 6%, G: 5%, H: 4%。给出一个最佳编码，使得通讯中出现的二进制数字尽可能少。

- 分析：用较短（长）的序列去表示出现频率高（低）的字母

- 问题转化为：
  - 求出叶的权分别为0.04，0.05，0.06，0.1，0.1，0.15, 0.20, 0.3的最优二叉树
  - 然后用该二叉树产生前缀编码传输上述给定的字母。

例：在计算机通信中要传输A, B, C, D, E, F, G, H 八个字母，它们出现频率为 A: 30%, B: 20%, C: 15%, D: 10%, E: 10%, F: 6%, G: 5%, H: 4%。给出一个最佳编码，使得通讯中出现的二进制数字尽可能少。

- **Huffman编码**
  1) 给定字母集$C=\{c_1, c_2,\ldots, c_n\}$及频率 $f(c_1),\ldots, f(c_n)$；
  2) 设有n个叶结点，分别以$f(c_1),\ldots, f(c_n)$为权；
  3) 在所有入度为0的结点，选出两个权最小的结点v, v'，添加一个新的分支节点u，使得u 以v和v'为儿子结点，且 $f(u)=f(v)+f(v')$；
  4) 重复3)直至只有一个入度为0的结点。

Huffman 算法

- 建立Huffman树的主要运算是插入和删除最小频率字符，所以用最小堆。

算法HUFFMAN

输入：n个字符的集合$C=\{c_1,c_2,\ldots, c_n\}$及频率 $\{ f(c_1), f(c_2),\ldots, f(c_n) \}$。

输出：C的Huffman树（V，T）。

1.根据频度将所有字符插入最小堆H

2.$V \leftarrow$C； T={}

3.for $j \leftarrow 1$ to n-1

4. $c \leftarrow$DELETEMIN(H)

5. $c' \leftarrow$DELETEMIN(H)

6. $f(v) \leftarrow f(c)+ f(c')$ //新节点v

7. INSERT (H, v)

8. $V = V \cup \{v\}$

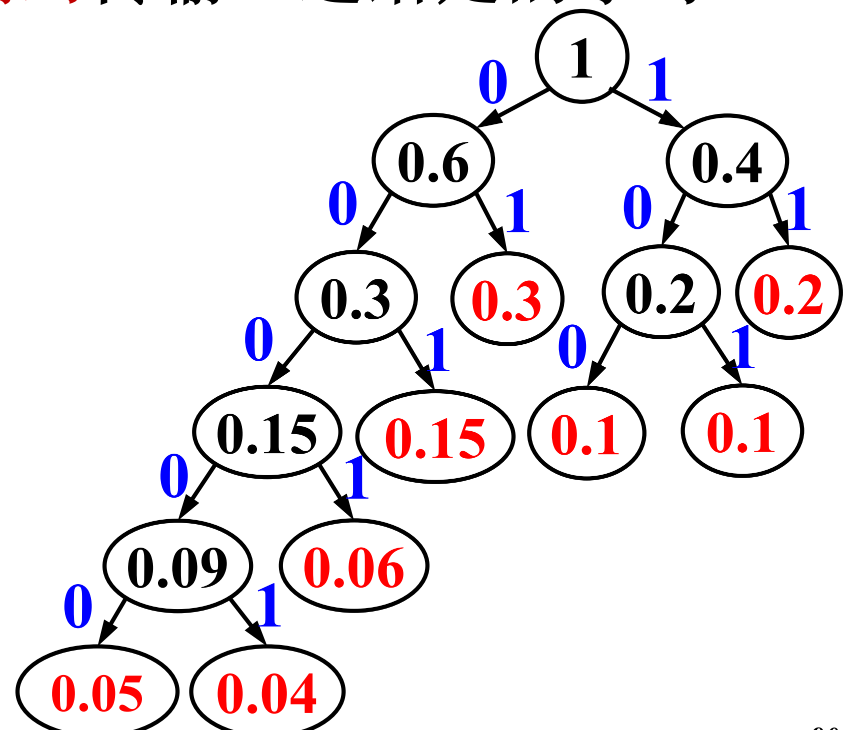
9. $T = T \cup \{(v, c), (v, c')\}$ // c, c'为T中v的孩子

10.end for

例：在计算机通信中要传输A, B, C, D, E, F, G, H 八个字母，它们出现频率为 A: 30%, B: 20%, C: 15%, D: 10%, E: 10%, F: 6%, G: 5%, H: 4%。给出一个最佳编码，使得通讯中出现的二进制数字尽可能少。

解：求出叶的权分别为0.04, 0.05, 0.06, 0.1, 0.1, 0.15, 0.20, 0.3的最优二叉树，然后用该二叉树产生前缀编码传输上述给定的字母。

0.04	0.05	0.06	0.1	0.1	0.15	0.20	0.3
	0.09	0.06	0.1	0.1	0.15	0.20	0.3
		0.15	0.1	0.1	0.15	0.20	0.3
		0.15		0.2	0.15	0.20	0.3
				0.2	0.3	0.20	0.3
					0.3	0.40	0.3
						0.40	0.6
							1.0

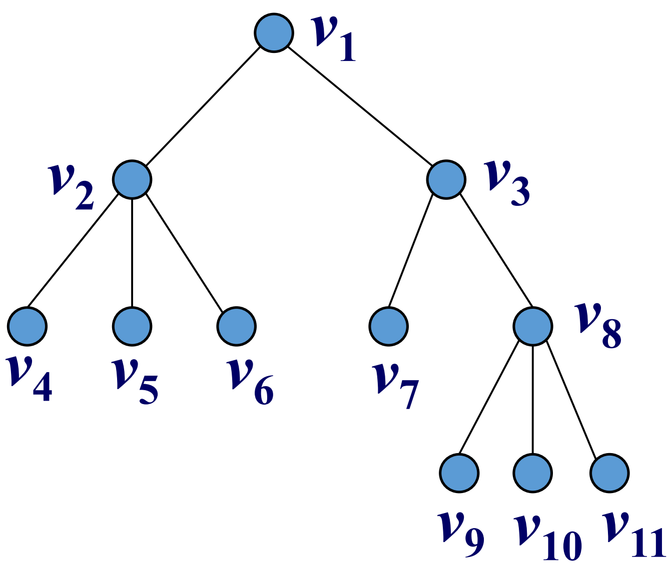


最佳编码

A: 01, B: 11, C: 001, D:100, E:101, F: 0001, G: 00000, H:00001

有序树

- 借用家族树的名称来称呼有序树的结点。



v_1 是 v_2 和 v_3 的父亲

v_2 是 v_1 的长子

v_2 是 v_3 的哥哥

v_6 是 v_5 的弟弟

v_2 是 v_7 的伯父

v_6 是 v_7 的堂兄

- 可以用定位二元有序树 表示有序森林。
- 有序森林和定位二元有序树之间建立一一对应关系。
  - 称位于左边的有序树之根为位于右边的有序树之根的哥哥

有序森林和定位有二元有序树之间的自然对应关系：

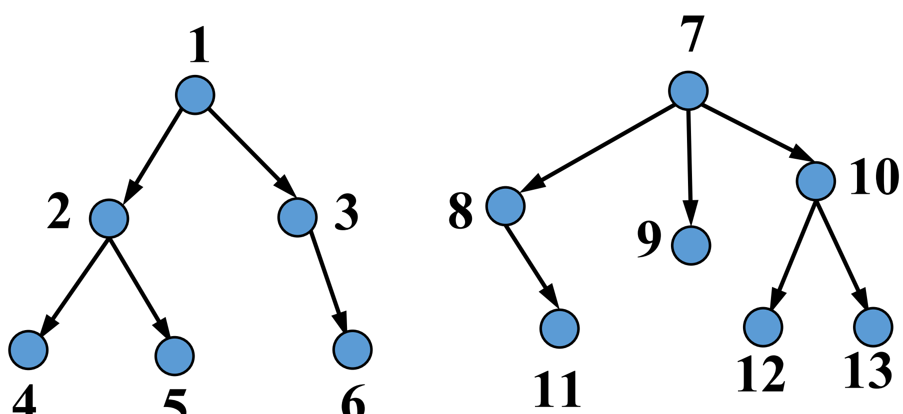
规定：*F* 与 *T* 有相同的结点。

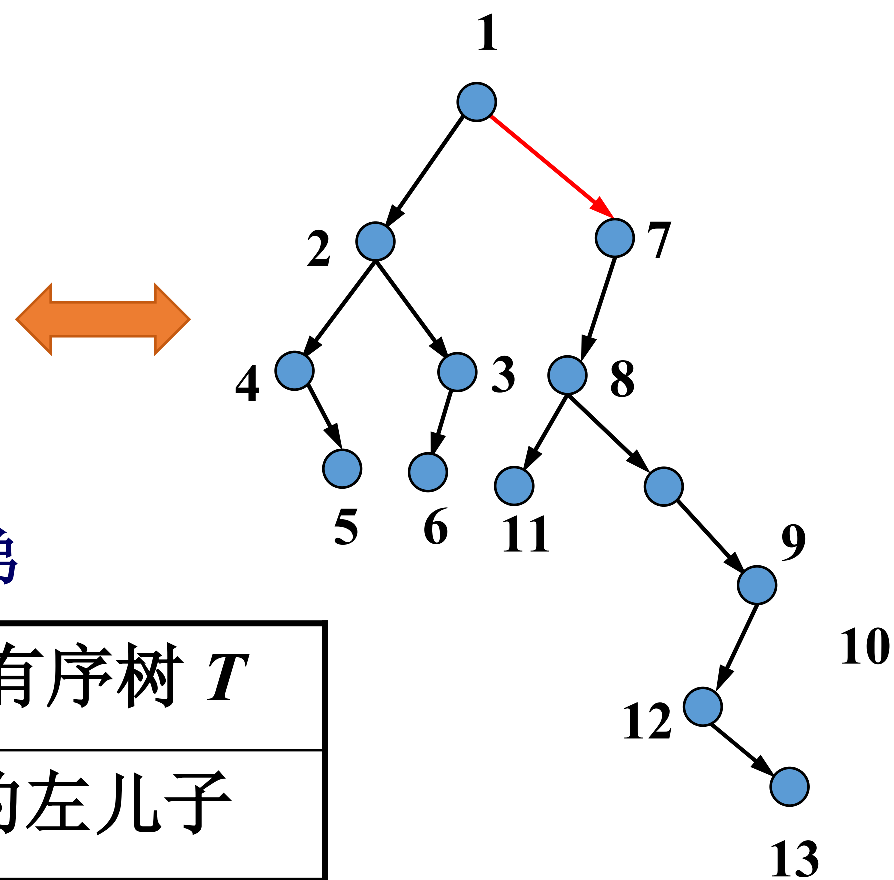
有序森林 *F*	定位二有序树 *T*
v_1是v_2的长子	v_1是v_2的左儿子
v_1是v_2的大弟	v_1是v_2的右儿子

亲弟弟，
不包括堂兄弟

定位有序树

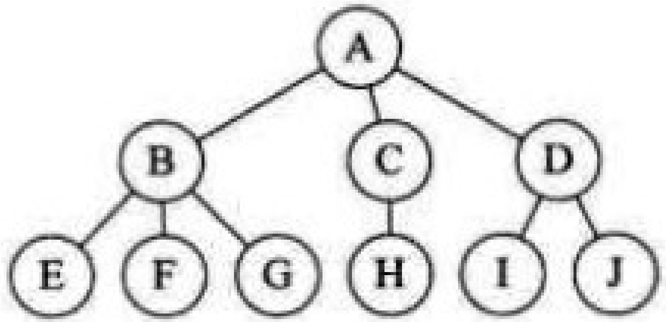
- 有序森林和定位二元有序树之间可以建立一一对应关系。



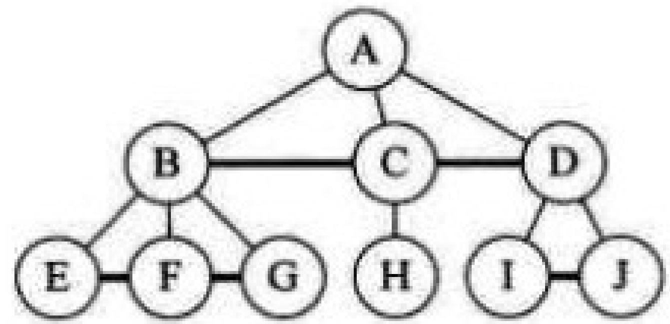


7 是1的大弟

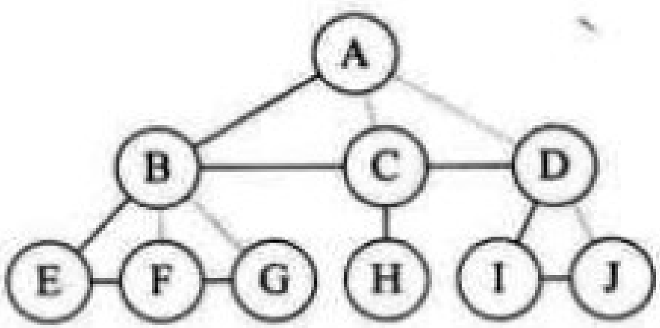
有序森林 *F*	定位二有序树 *T*
v_1是v_2的长子	v_1是v_2的左儿子
v_1是v_2的大弟	v_1是v_2的右儿子



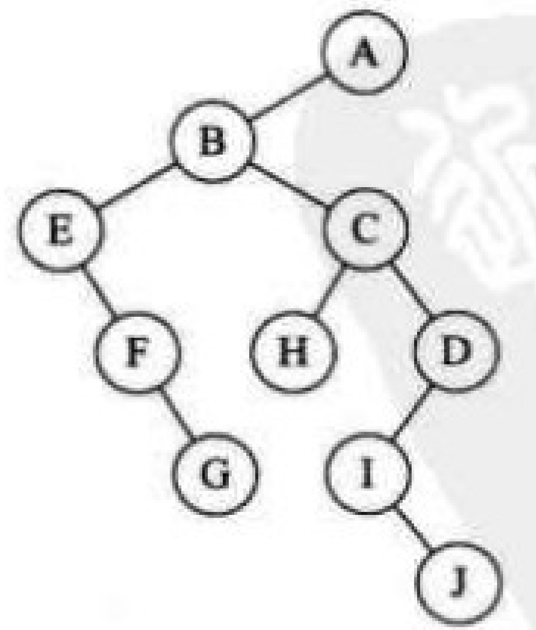
树



步骤1：给兄弟加线



步骤2：给除长子外的孩子去线



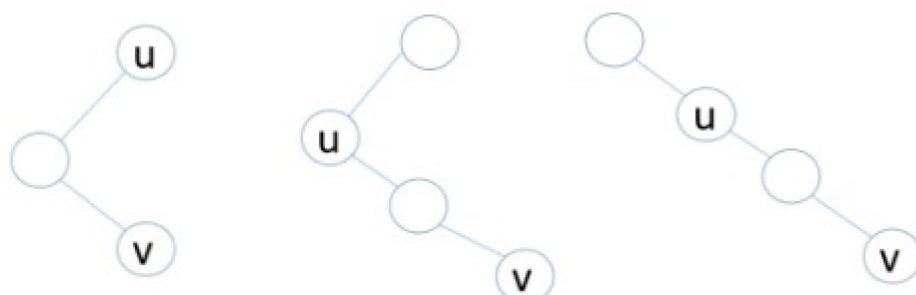
步骤3：层次调整

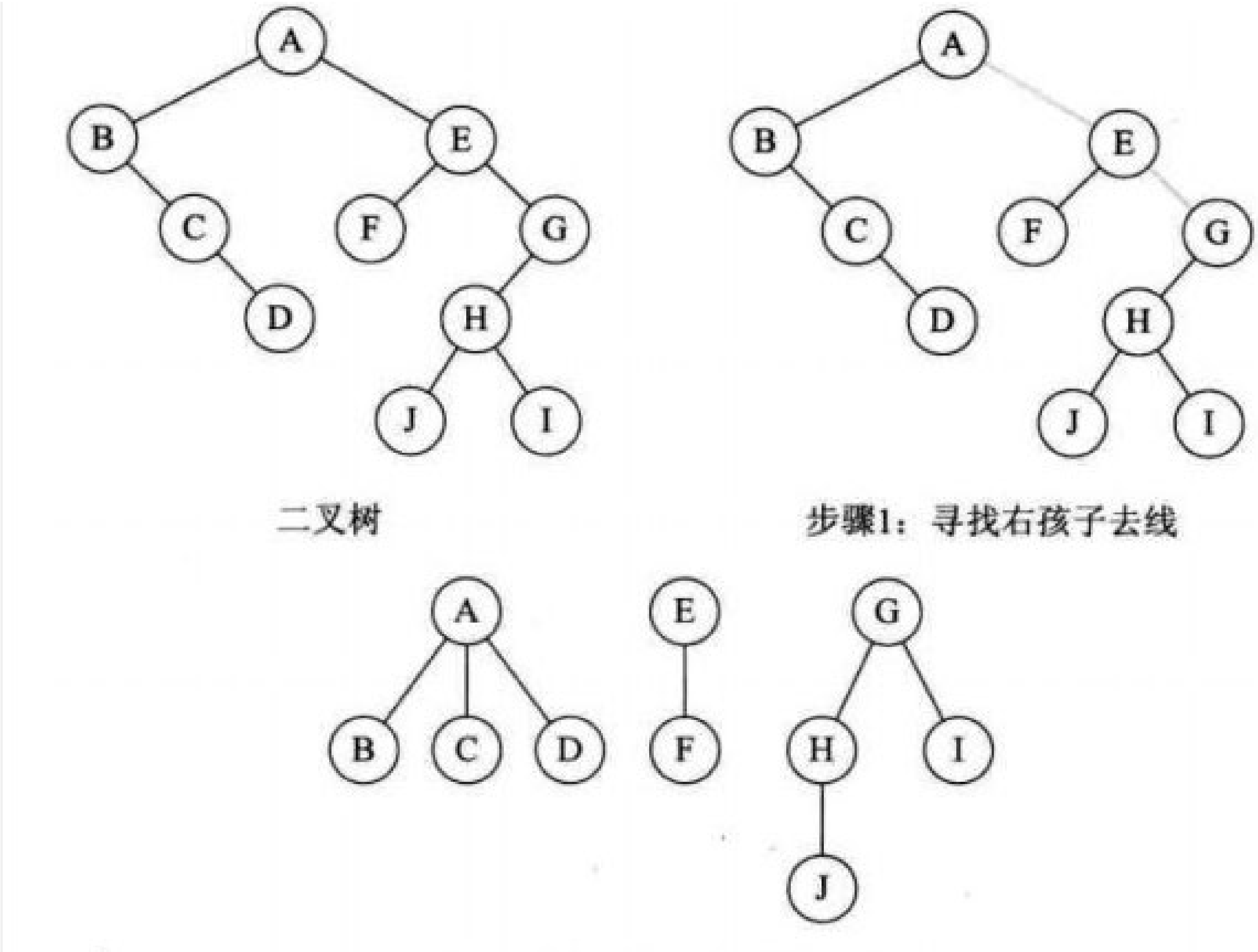
将森林转换为对应的二元有序树，若在二元有序树中，结点 *u* 是结点 *v* 的父结点的父结点，则在原来的森林中，*u* 和 *v* 可能具有的关系是（ ）

I. 父子关系

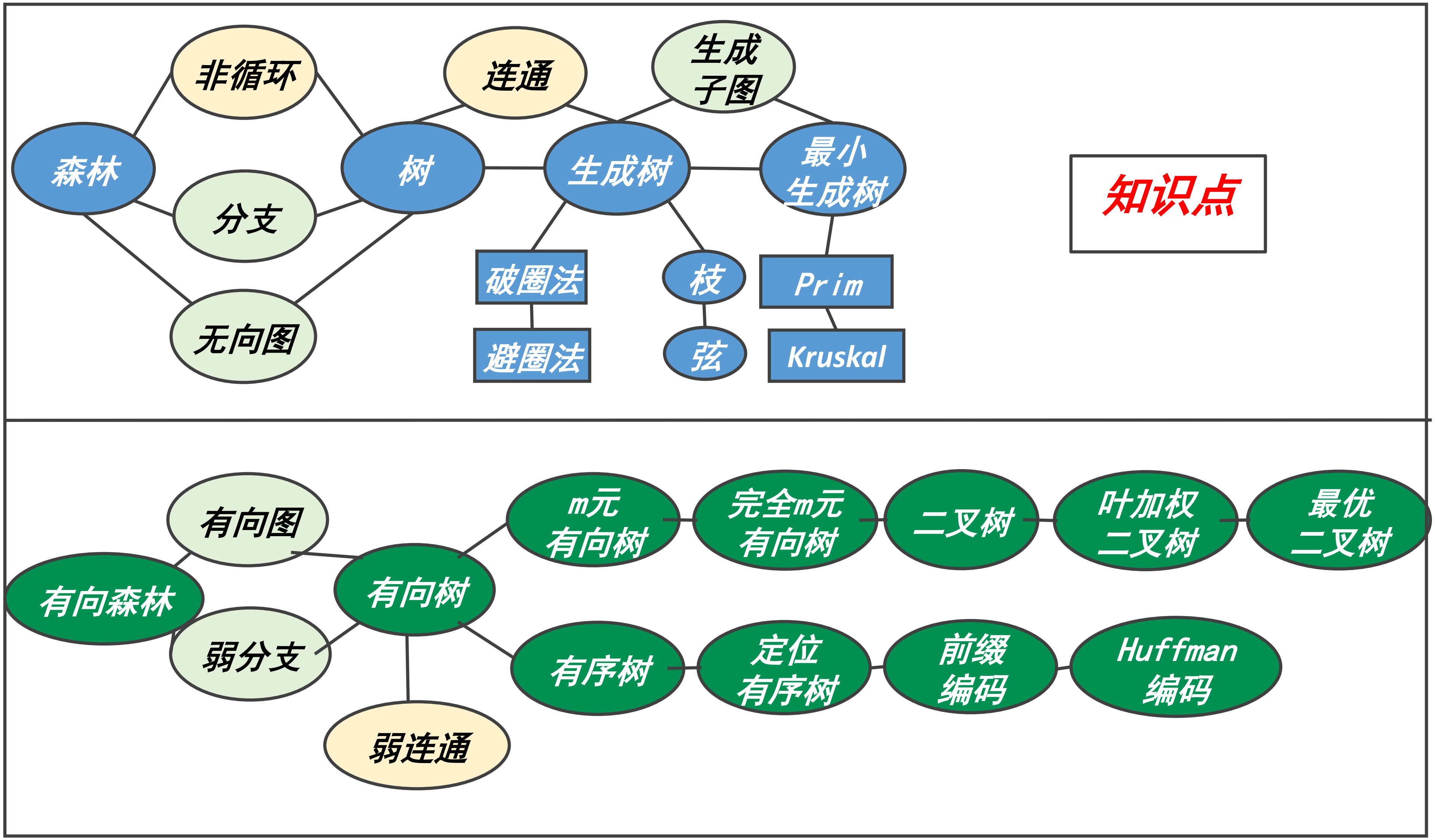
II.兄弟关系

III. u 的父结点与 **v** 的父结点是兄弟关系




二叉树
步骤1：寻找右孩子去线

本章知识点


非循环
连通
生成
子图
森林
树
生成树
最小
生成树
分支
破圈法
枝
Prim
避圈法
弦
Kruskal
无向图
知识点
有向图
m元
有向树
完全m元
有向树
二叉树
叶加权
二叉树
最优
二叉树
有向森林
有向树
弱分支
有序树
定位
有序树
前缀
编码
Huffman
编码
弱连通