L7：树与二叉树基础

吉林大学计算机学院
谷方明
fmgu2002@sina.com

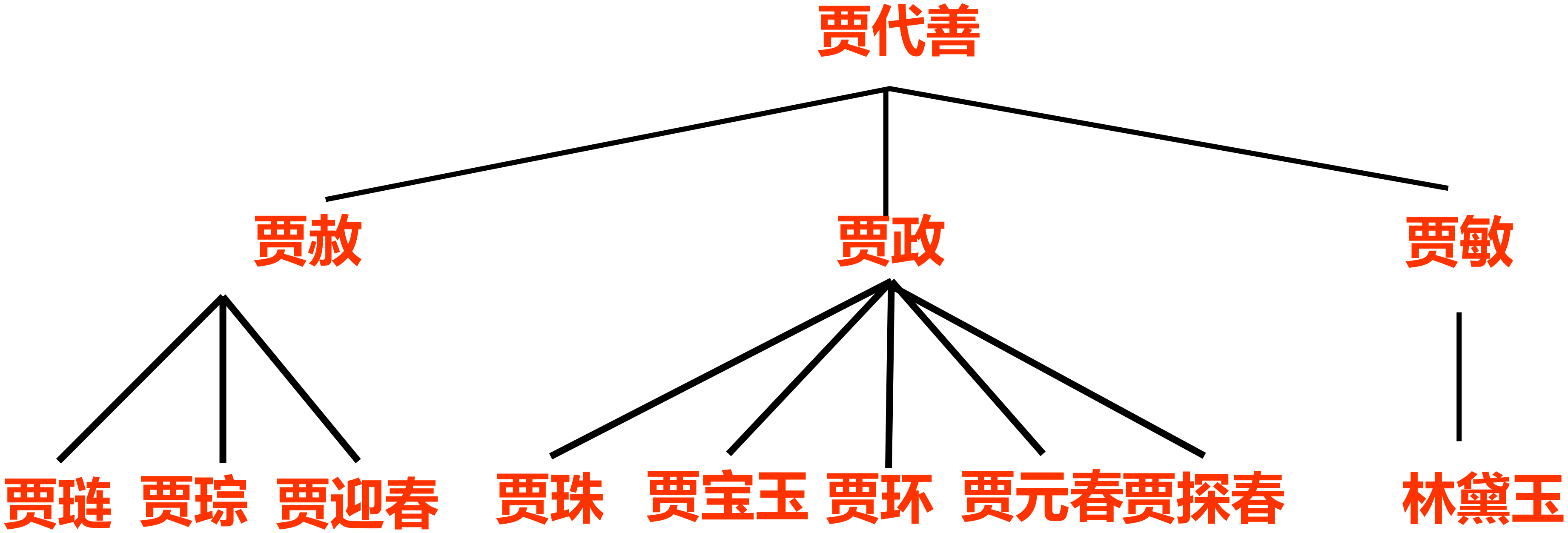
学习目标

- 掌握树的定义、术语和表示
- 掌握二叉树、完全二叉树、满二叉树的定义
- 掌握二叉树的五个基本定理

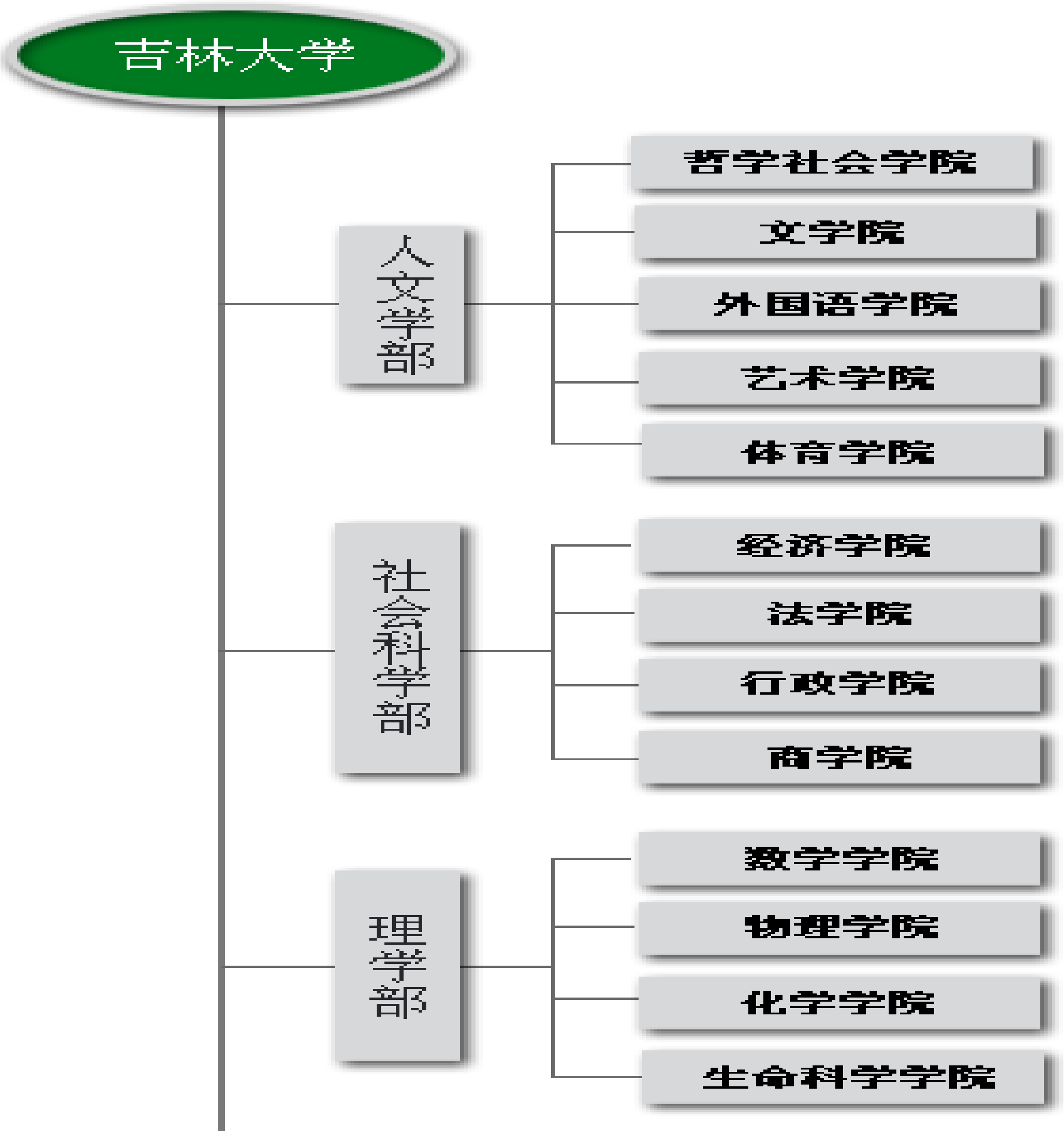
5.1 树的基本概念

- 树形结构是一种非常重要的非线性数据结构。可用来描述客观世界中广泛存在的具有分支或层次关系的对象。

例1 族谱(直系图)



例2 组织结构



吉林大学
人文学部
哲学社会学院
文学院
外国语学院
艺术学院
体育学院
社会科学部
经济学院
法学院
行政学院
商学院
理学部
数学学院
物理学院
化学学院
生命科学学院

例3 计算机领域的广泛应用

菜单 1

2 文件系统

程序 3

4 ……

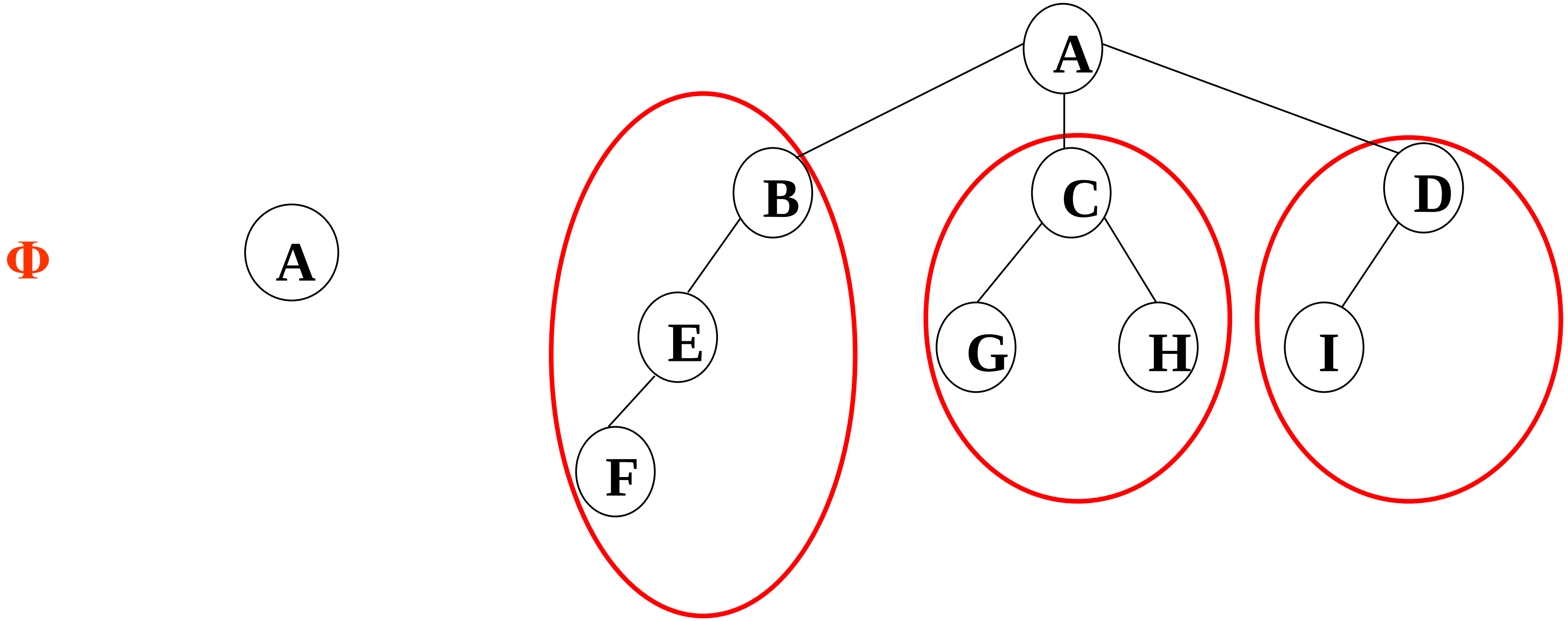
树的定义

定义5.1： 一棵树是一个有限的结点集合T. 若T空，则称为空树。若T非空，则：

1. 有一个被称为根的结点，记为$root(T)$；
2. 其余结点被分成$m(m \geq 0)$个不相交的非空集合T_1，T_2，…，T_m，且T_1，T_2，…，T_m又都是树。

树T_1，T_2，…，T_m称作$root(T)$的子树。

例4



(a)空树　　(b)单个结点的树　　(c)多个结点的树

有序树

- 如果树的定义(2)中，子树$T_1, T_2, \dots, T_m$ 的相对次序是重要的，则称该树为有序树，否则称为无序树。在有序树中，把T_i 称作根的第 i 个子树。

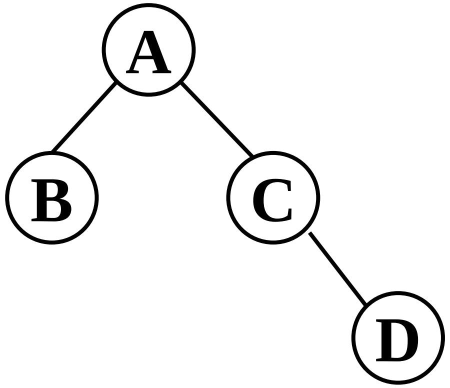


图5.2 （a）树1

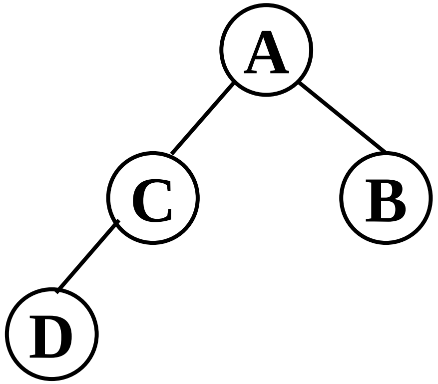
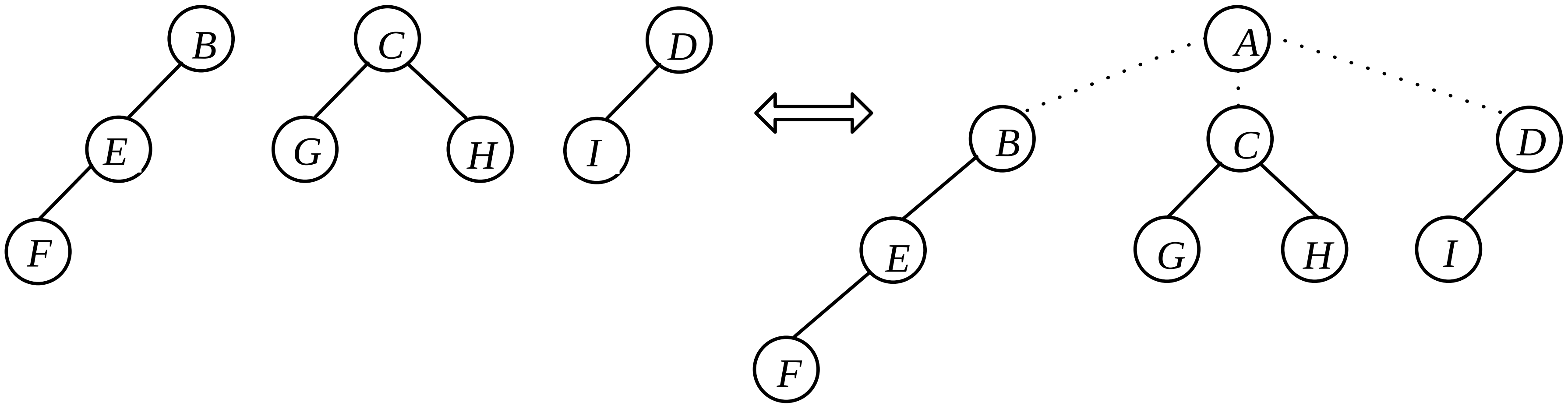


图5.2 （b）树2

- 因为计算机表示定义了树的一种隐含次序，所以假定所讨论的树都是有序的.

森林

- 一个森林是若干个非空不相交的树的集合
- 抽象的森林和树的差别很小。如果删掉一个树的根，就得到一个森林；反过来，增加一个结点作为根结点到任何森林，就得到一个树。树和森林这两个词几乎可互换的使用。



树的定义说明

- 本节定义的树在图论中称为有根树
  - 有根树(rooted tree)：要求选出一个结点特别标识，一般称为根(root)。
  - 无根树(unrooted tree)：连通无环图，也叫自由树(free tree)
- 数据结构的树：有限的带标号的有根有序的树（图论术语）

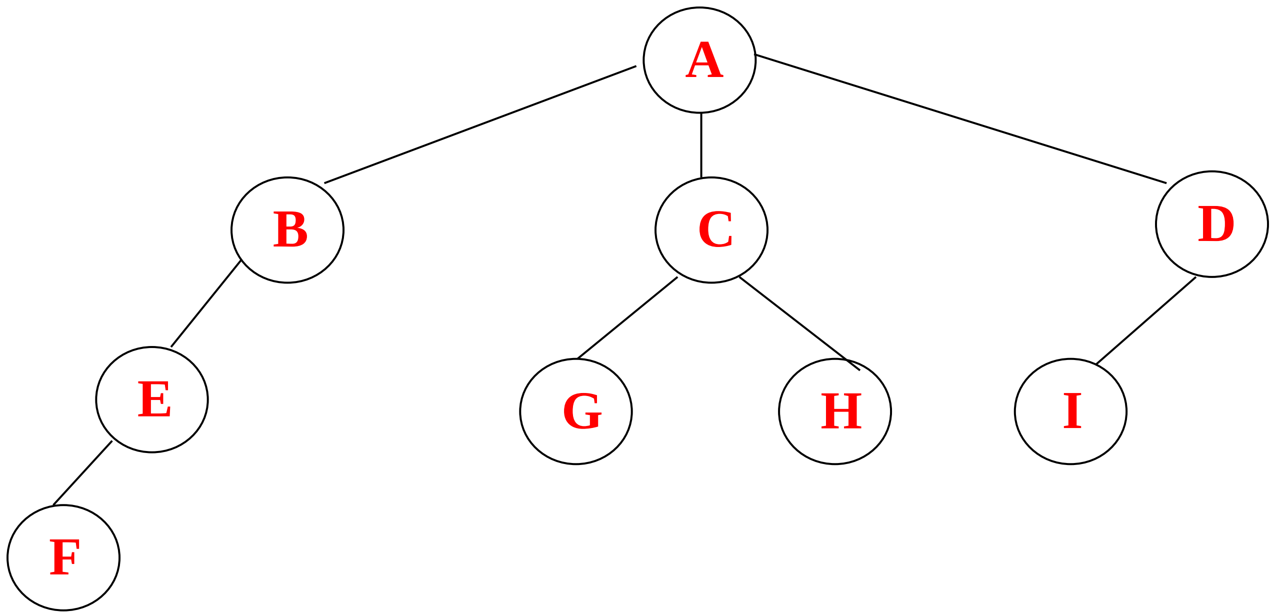
树的术语 I
父亲、儿子、兄弟、祖先、后裔

1. 父亲、儿子、兄弟：

   每个结点都是它的子树的根的父亲；反过来，每个结点都是它父亲的儿子；同一父亲的儿子是兄弟。

2. 祖先、后裔：

   每个结点都是它的子树的所有结点的祖先；反过来，每个结点都是它祖先的后裔。


A
B
C
D
E
G
H
I
F

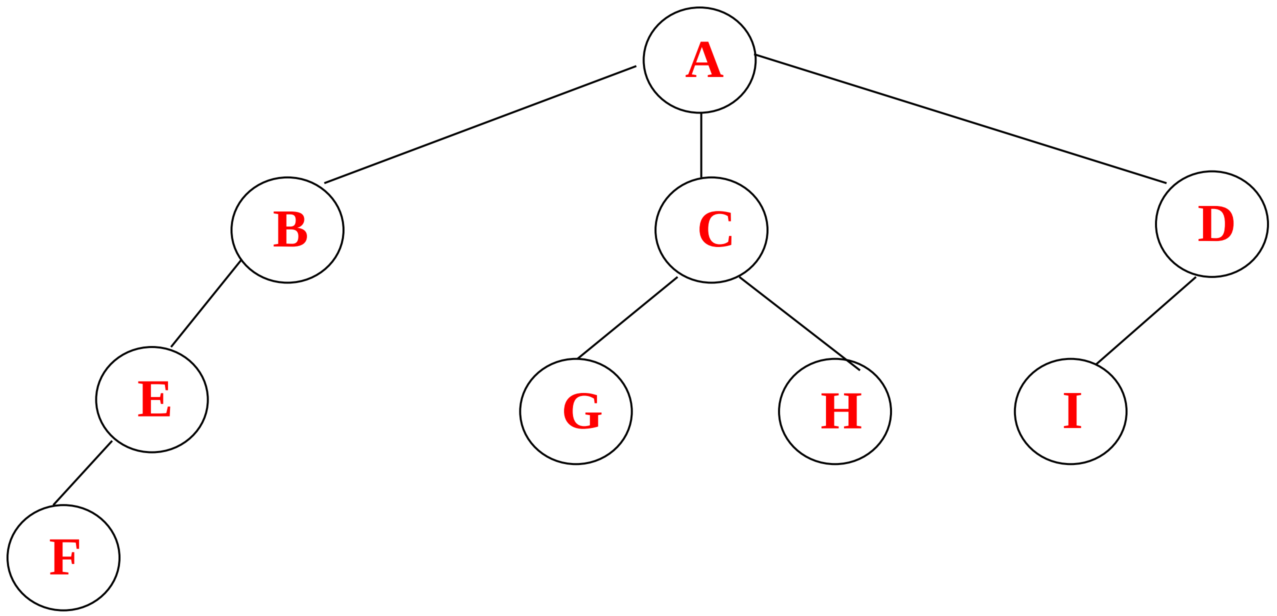
树的术语 II
度、叶子、分枝

1. **度**

   一个结点的子树的个数，称为该结点的度或次数。一棵树的度为 $\max_{i=1,\ldots,n} D(i)$.

2. **叶结点、分支结点**：

   度为0的结点称为终端结点；非根的终端结点称为叶结点，有时不严格区分终端结点和叶结点；非终端结点称为分枝结点。

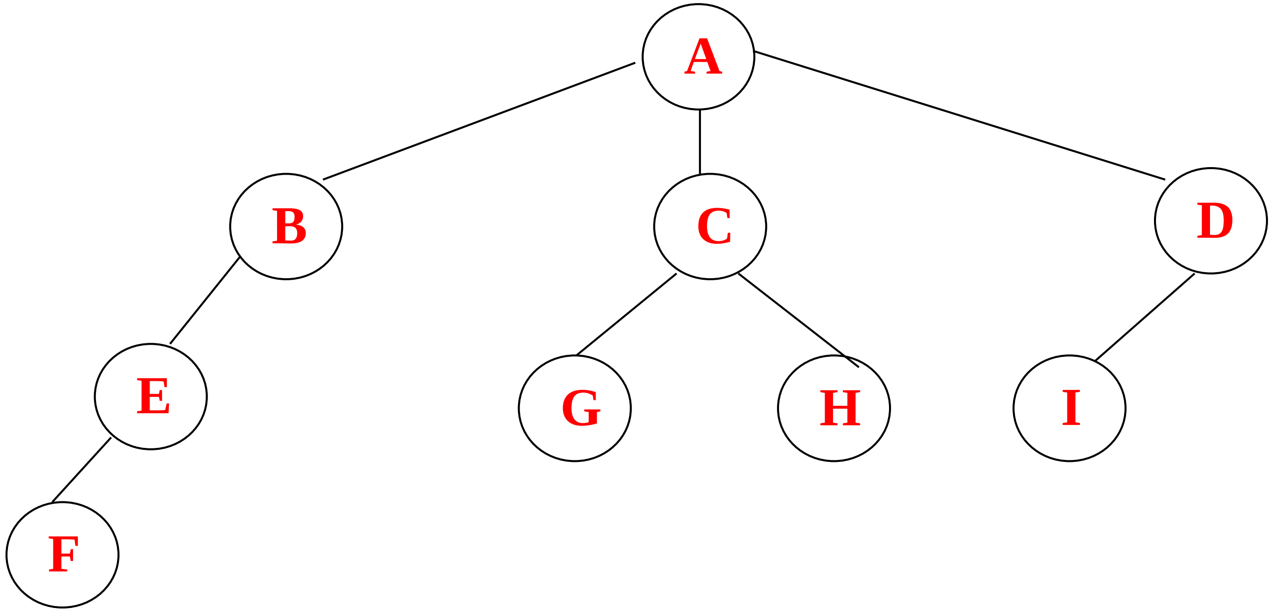

A
B
C
D
E
G
H
I
F

树的术语 III
结点的层数

1. <u>结点的层数</u>

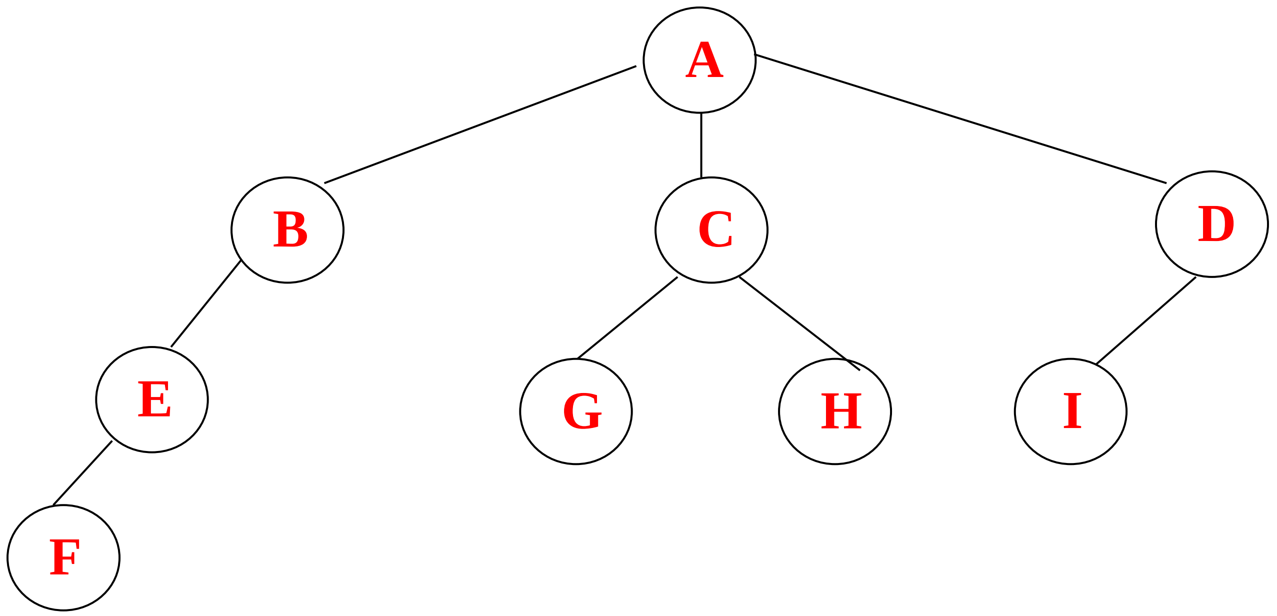
   ⑴ ***root*(*T*)**层数为零；

   ⑵ 其余结点的层数为其父结点的层数加 **1**；



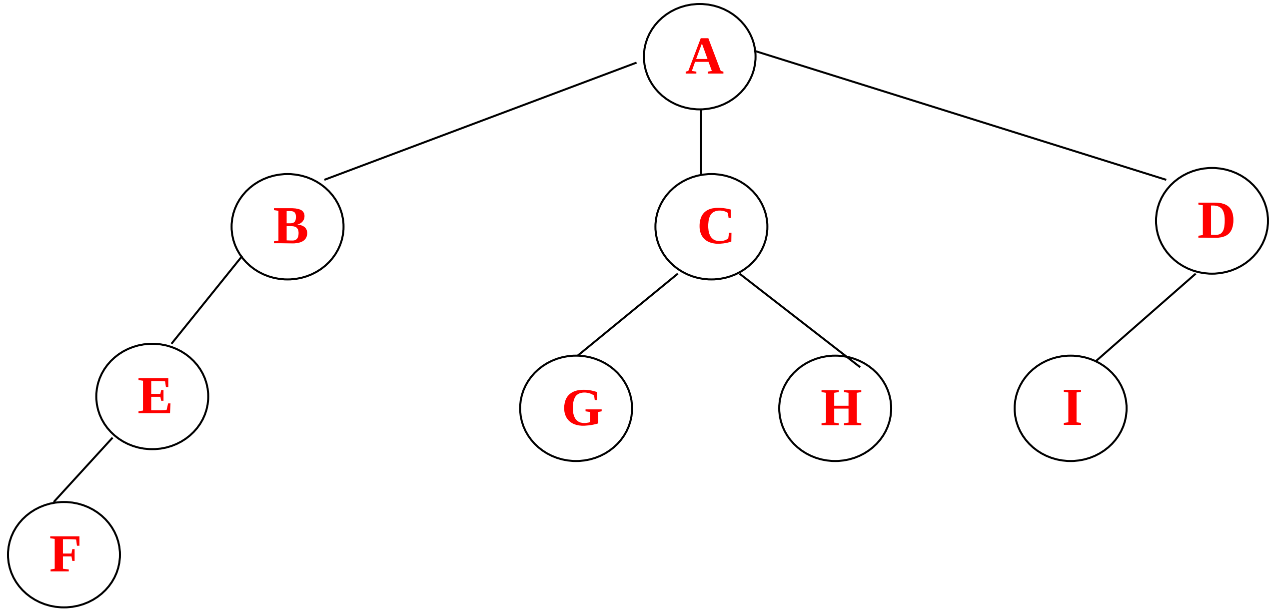
树的术语 IV
路径、深度、高度

- 结点序列$v_1, v_2, \ldots, v_k$称为结点v_1到结点v_k的<u>路径</u>，如果v_i是v_{i+1}的父亲$(1 \le i < k)$. 路径经历的边数称为<u>路径长度</u>，即$k-1$.
- 根到结点v_i的路径长度称为结点v_i的<u>深度</u>。设n为树的结点数，$D(i)$表树中第i个结点的深度，则一棵 <u>树的深度</u>为 $\max_{i=1,\ldots,n} D(i)$
- 结点v_i到叶结点的最长路径的长度称为结点v_i的<u>高度</u>。叶结点的高度为0；一棵<u>树的高度</u>定义为其根结点的高度。

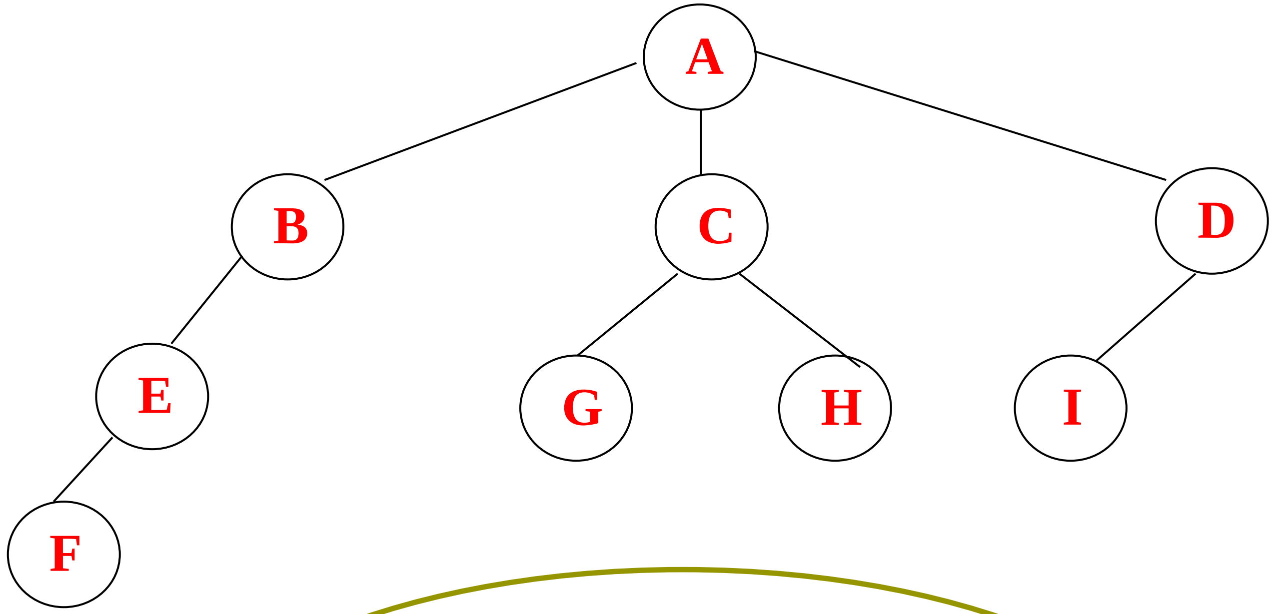

A
B
C
D
E
G
H
I
F

树的表示法

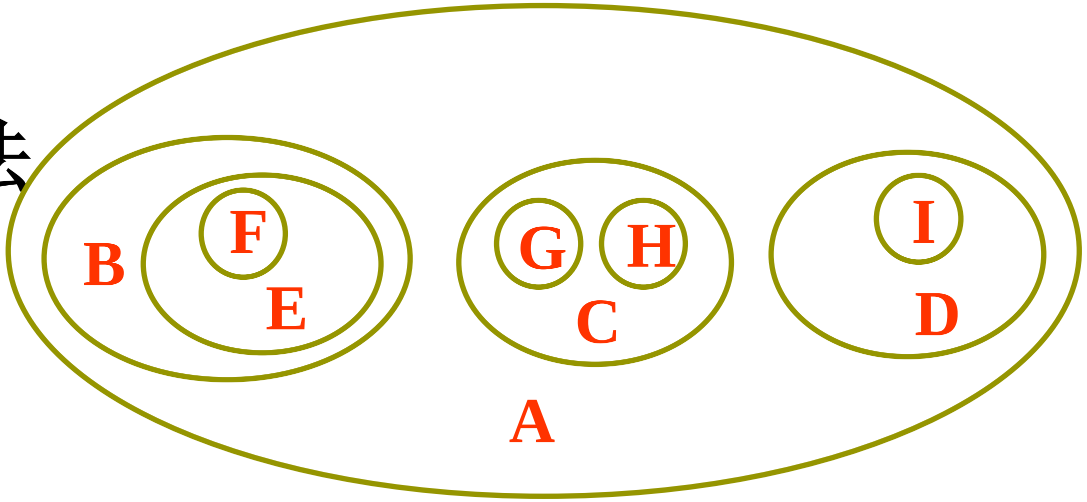
- 树形表示法
  - ✓ 树根的位置



树的表示法

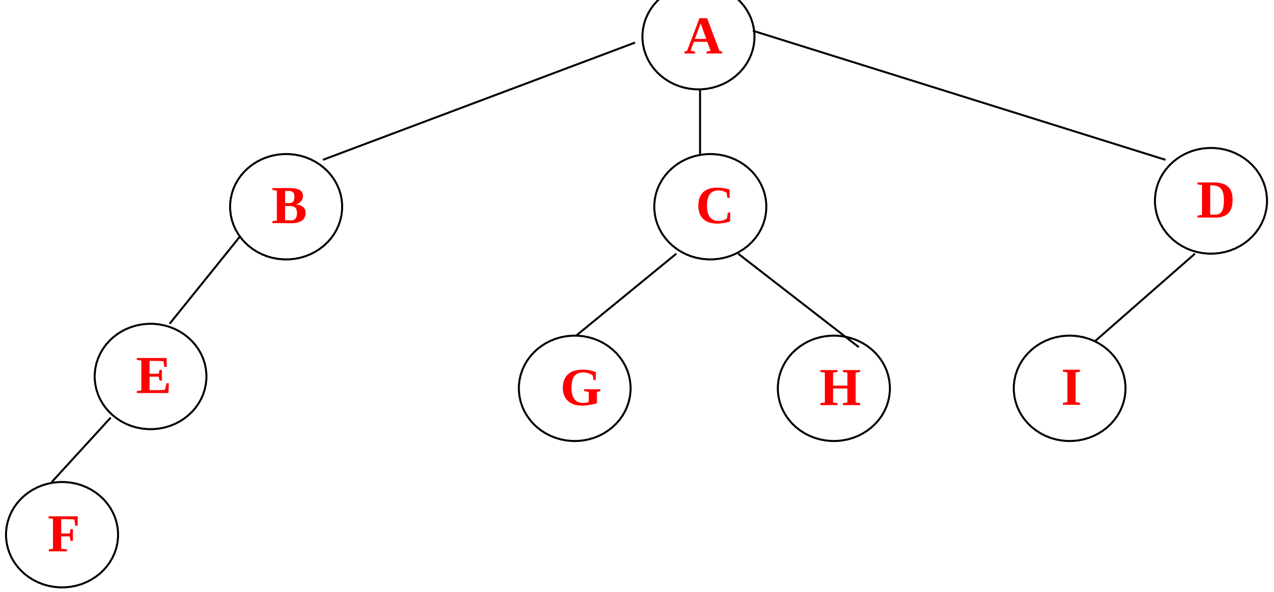


- 嵌套集合表示法



{ A,{B,{E,{F}}}, {C,{G},{H}},{D,{I}} }

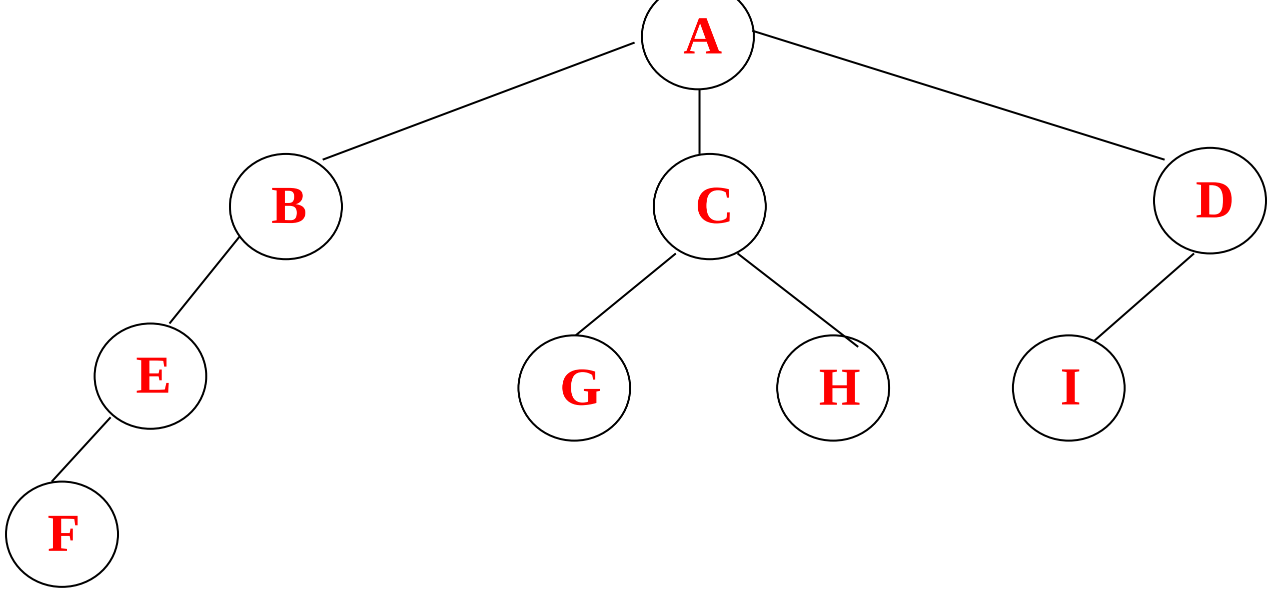
树的表示法



{ A,{B,{E,{F}}}, {C,{G},{H}},{D,{I}} }

- 嵌套括号表示法
  - ✓ (A (B(E(F)), C(G,H), D(I)))

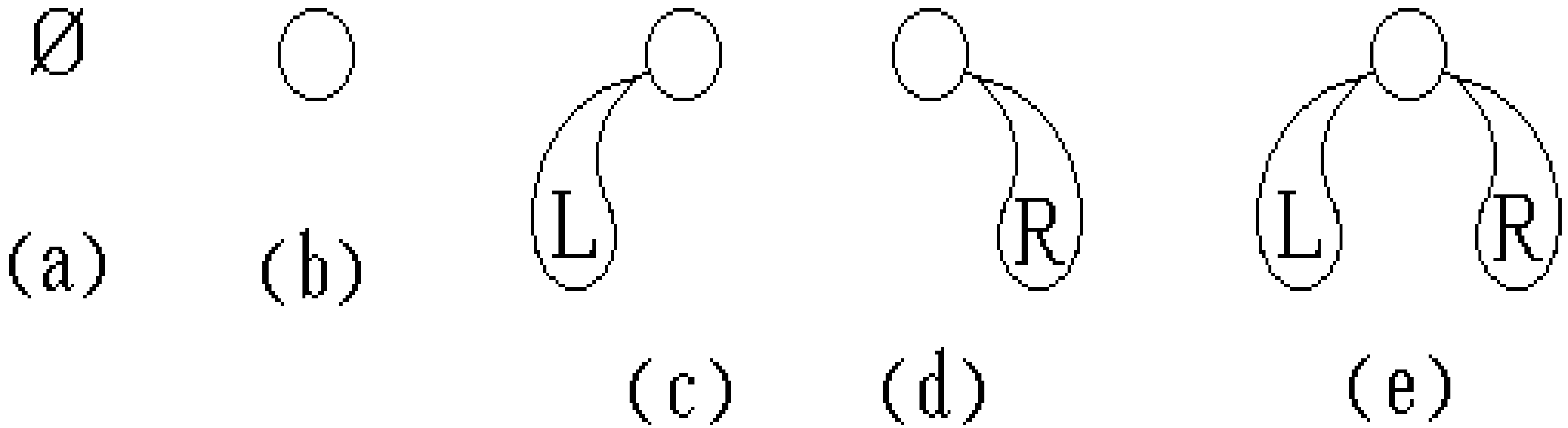
树的表示法



- 凹入表示法

5.2 二叉树

- 定义5.2：二叉树是结点的有限集合，它或者是空集，或者由一个根及左、右两个不相交的的二叉树构成。这两棵子树分别称为左、右子树。

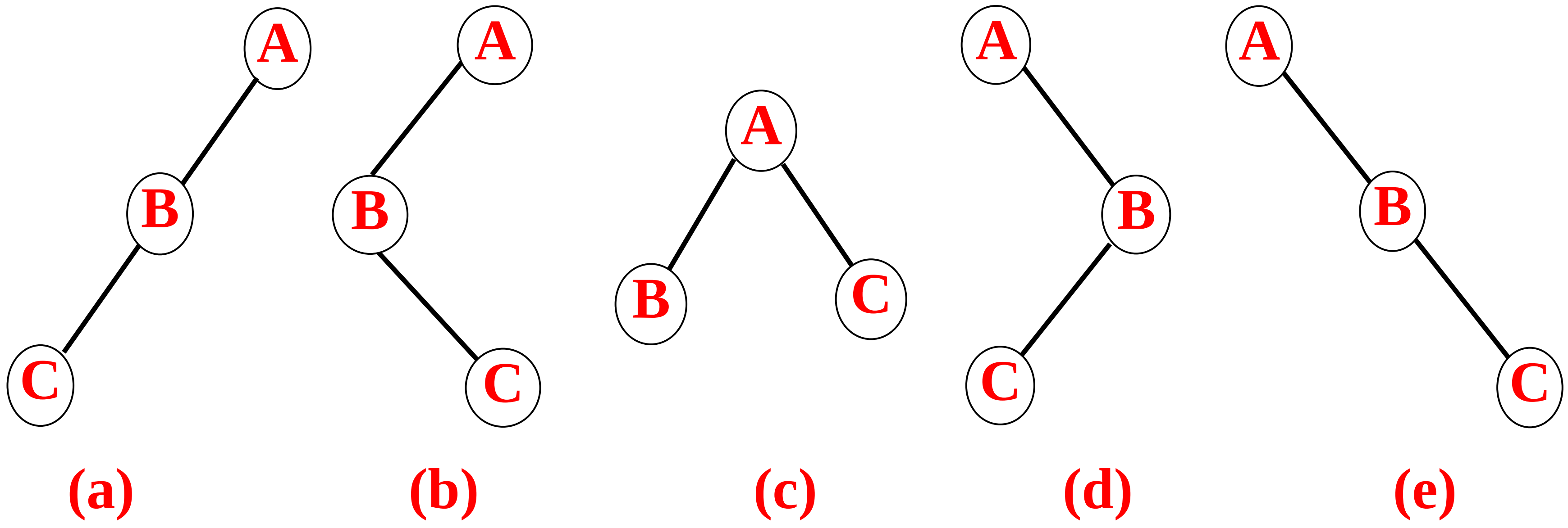


(a) (b) (c) (d) (e)

树与二叉树的主要区别

I. 二叉树每个结点最多有 2 个子结点，树则无此限制；
II. 二叉树中结点的子树分成左子树和右子树，即使某结点只有一棵子树，也要指明该子树是左子树，还是右子树，即二叉树不仅是有序的，而且是拓扑分叉的

例 含有3个结点的不同二叉树



含有3个结点的不同二叉树

引理5.1

二叉树中层数为i的结点**至多**有2^i个，$i \geq 0$。

引理5.2

高度为k的二叉树中至多有$2^{k+1}-1$ ($k \geq 0$)个结点。

引理5.3

- 设T是由n个结点构成的二叉树，其中叶子结点个数为n_0, 度为2的结点个数为n_2，则有：$n_0 = n_2 + 1$

证明：设度为1的结点有n_1个，总结点个数为n，总边数为e，则：

$n=n_0+n_1+n_2$　(1)

$e=n-1$　(2)

$e=2n_2+n_1$　(3)

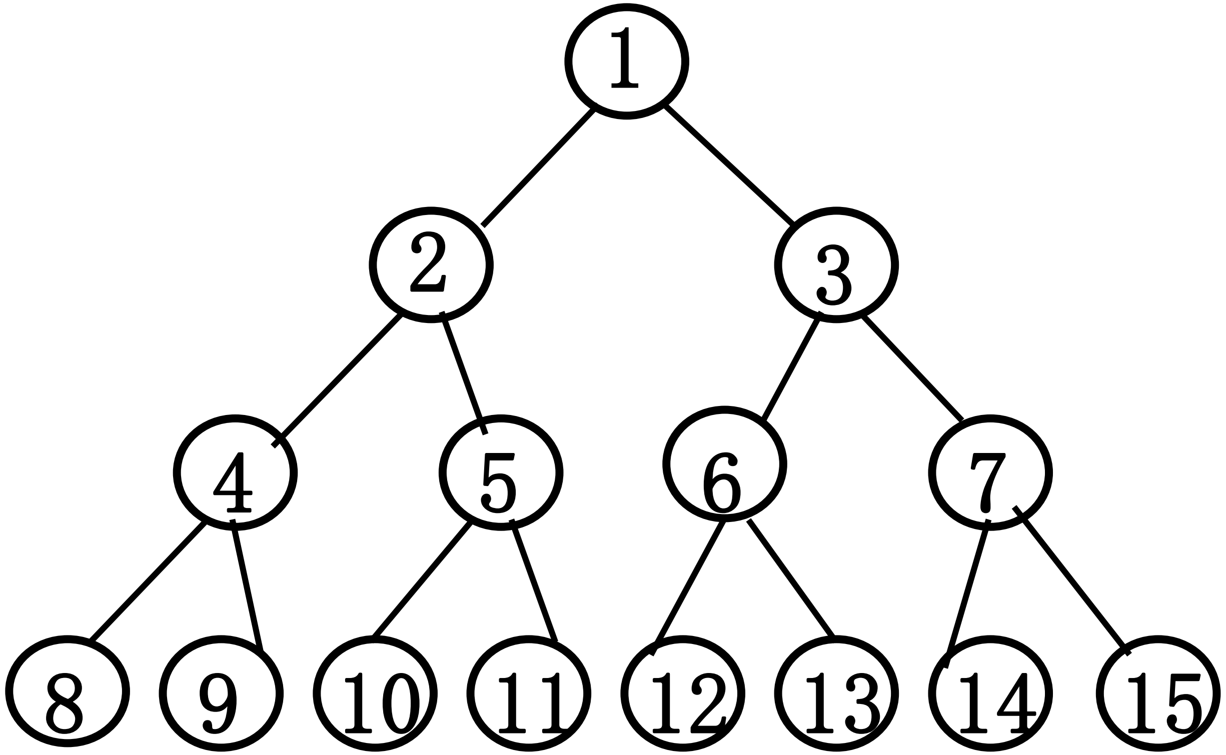
因此，有$n_0+n_1+n_2-1=2n_2+n_1$

$n_0=n_2+1$.

证毕■

满二叉树

- 定义5.3 一棵非空高度为$k(k \ge 0)$的<u>满二叉树</u>，是有$2^{k+1}-1$个结点的二叉树。



满二叉树的特点

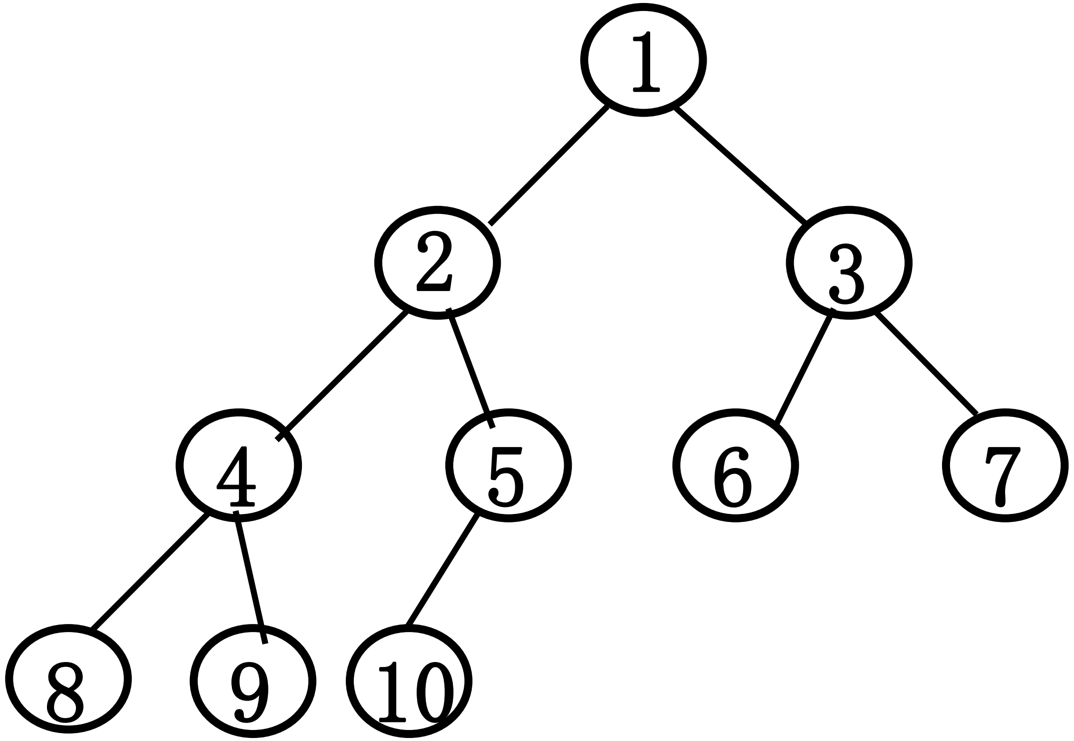
① 叶结点都在第*k*层上；

② 每个分支结点都有两个子结点；

③ 叶结点的个数等于非叶结点个数加1 ；

完全二叉树

- **定义5.4** 一棵包含n个结点高为k的二叉树T，当按层次顺序编号T的所有结点，对应于一棵高为k的满二叉树中编号由1至n的那些结点时，T被称为完全二叉树。



完全二叉树的特点

- 树中叶结点只能在层数最大的两层上出现；树中最下一层的结点都集中在该层最左边的若干位置上；
- 仅仅编号最大的分支结点可以没有右孩子，其余分支结点都有两个孩子；

引理5.4

- 若将一棵具有n个结点的完全二叉树按层次次序从1开始编号，则对编号为i ($1 \le i \le n$)的结点有：

① 若$i \ne 1$，则编号为i的结点的父结点的编号为$\lfloor i/2 \rfloor$。

② 若$2i \le n$，则编号为i的结点的左孩子的编号为 $2i$，否则 i 无左孩子。

③ 若$2i+1 \le n$，则i结点的右孩子结点编号为$2i+1$，否则 i 无右孩子。

引理5.5

- 具有n个结点的完全二叉树的高度是$\lfloor \log_2 n \rfloor$

小结

- 树的定义
- 树的术语（4组）
- 树的表示（4种）
- 二叉树的定义
- 二叉树的5个基本定理
- 完全二叉树和满二叉树