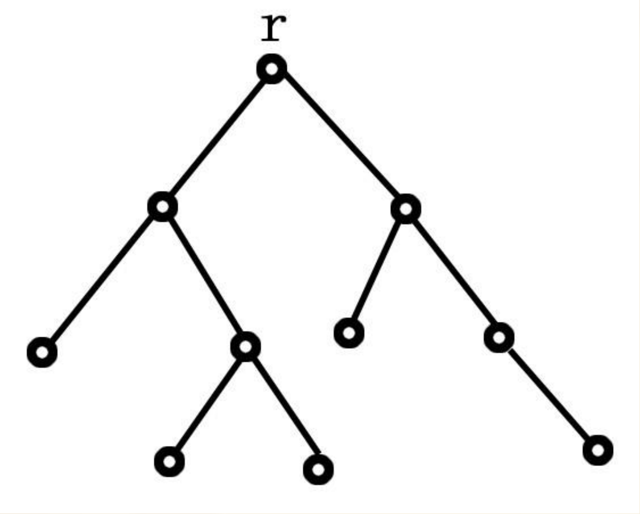
§16.5 裂项及其应用

定义 设T是一个有向图，若T的基础图是树，则称T为**有向树**。若一棵非平凡有向树T中有一个入度为0的顶点，其余顶点的入度均为1，则称T为**根树**。

在根树中，入度为0的顶点称为**树根**，入度为1且出度不为0的顶点称为**内点**，树根与内点统称为**分支点**；入度为1且出度为0的顶点称为**树叶**；从树根到任意顶点v的有向路径的长称为v的**层数**，所有顶点中最大层数称为**树高**。

根树的画法：



定义 设T是一棵非平凡根树，$\forall v_i, v_j \in V(T)$，若$v_i$可达$v_j$，则称$v_i$是$v_j$的祖先，$v_j$是$v_i$的后代；若$v_i$邻接到$v_j$，则称$v_i$是$v_j$的父亲，$v_j$是$v_i$的儿子；若$v_j$与$v_k$的父亲相同，则称$v_j$与$v_k$是兄弟。

有序树：同层顶点标定了次序的根树

***r*叉树**：每个分支点至多有r个儿子

***r*叉有序树**：有序的r叉树

***r*叉正则树**：每个分支点恰有r个儿子

***r*叉正则有序树**：有序的r叉正则树

***r*叉完全正则树**：r叉正则树，且所有树叶的层数均等于树高

***r*叉完全正则有序树**：有序的r叉完全正则树

定义 设T是一棵根树，v是T的一个分支点。由v及其后代导出的子图T_v称为T的以v为根的**根子树**。2叉正则树中一个分支点的两个儿子导出的根子树分别称为**左子树**和**右子树**。

定义 设T是一棵2叉树，有t片树叶$v_1,v_2,\cdots,v_t$，权分别为$w_1,w_2,\cdots,w_t$，则 $W(T)=\sum_{i=1}^{t} w_i l(v_i)$称为$T$的**权**，其中$l(v_i)$为$v_i$的层数。

给定树叶的权$w_1,w_2,\cdots,w_t$，则在全部有t片树叶的2叉树中，权最小者称为**最优2叉树**。

Huffman算法:

给定实数$w_1,w_2,\cdots,w_t$，$w_1\le w_2\le\cdots\le w_t$：

（1）连接权为w_1,w_2的两片树叶，得分支点，权为w_1+w_2;

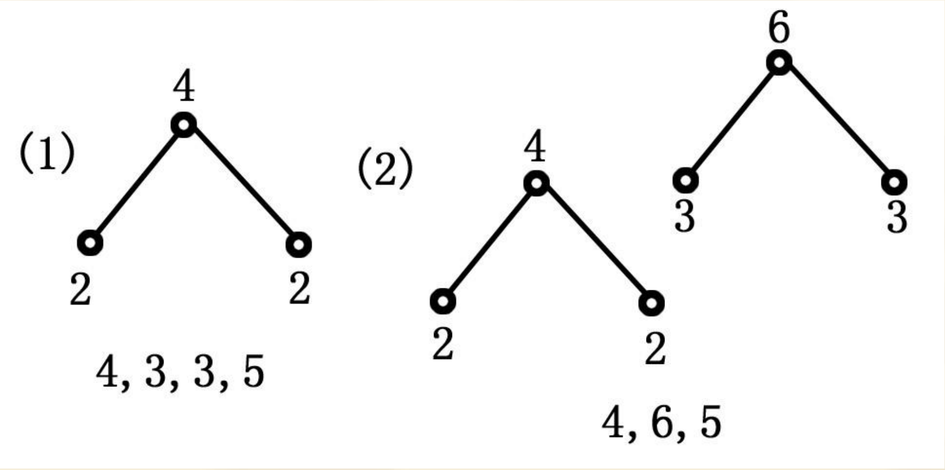
（2）在$w_1+w_2,w_3,\cdots,w_t$中选最小的两个权，连接它们对应的顶点，得新分支点及所带的权;

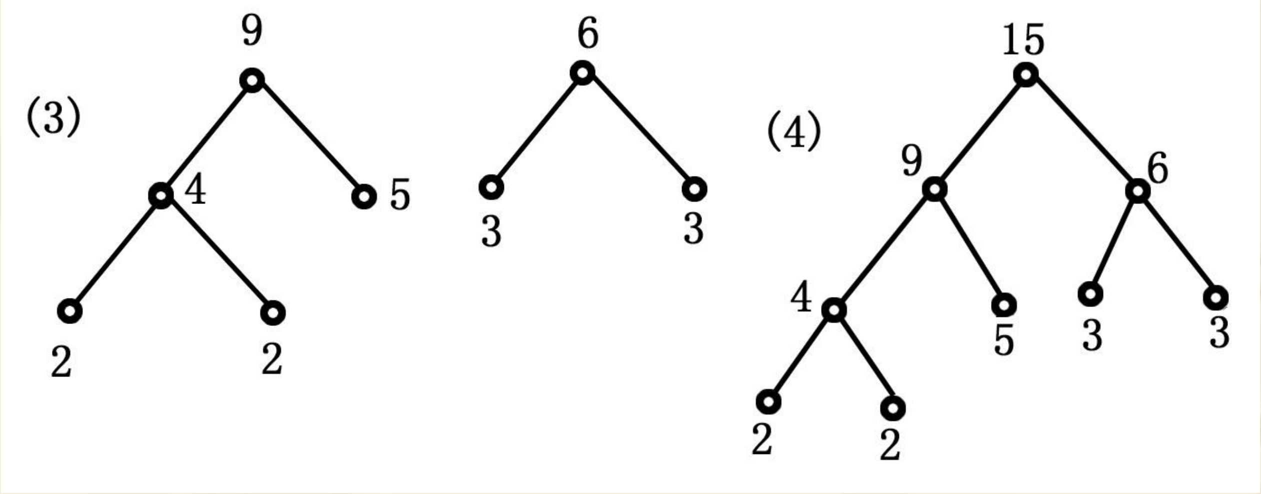
（3）重复(2)，直到形成$t-1$个分支点，t片树叶。

由此得到的2叉树权最小。

例 求带权2,2,3,3,5的最优2叉树。

解





2叉树的两个应用：

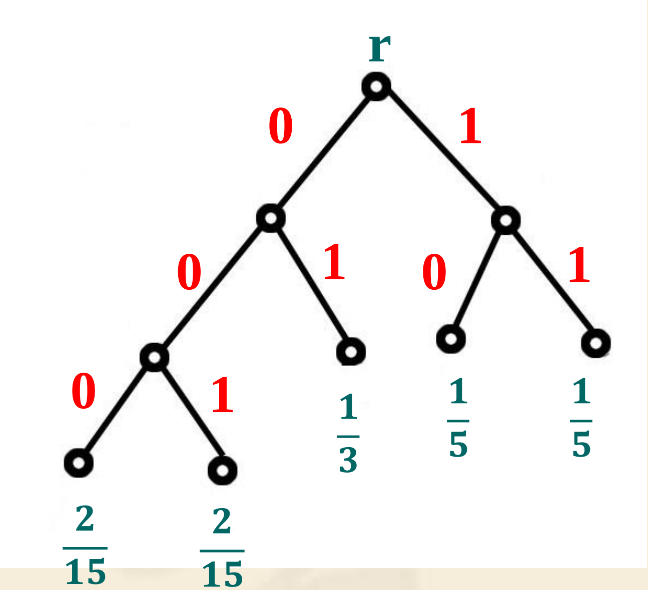
1. 最佳编码

定义 设$\alpha_1\alpha_2\cdots\alpha_n$是长为$n$的符号串，则称下列子串

$$\alpha_1,\ \alpha_1\alpha_2,\ \alpha_1\alpha_2\alpha_3,\ \cdots,\ \alpha_1\alpha_2\cdots\alpha_{n-1}$$

为该符号串的**前缀**；设$A=\{\beta_1, \beta_2, \cdots, \beta_m\}$是一个符号串的集合，若其中任意$\beta_i$与$\beta_j$ $(i\neq j)$互不为前缀，则称A为**前缀码**。由0,1符号串构成的前缀码称为**2元前缀码**。

例 设某系统需要传输5个基本符号，分别记为a,b,c,d,e，它们的出现频率分别为$\frac{2}{15},\frac{2}{15},\frac{1}{5},\frac{1}{5},\frac{1}{3}$。为这些符号构造2元前缀码。



解 以$\frac{2}{15},\frac{2}{15},\frac{1}{5},\frac{1}{5},\frac{1}{3}$为权构造最优2叉树，如上图。

现把每个分支点向下的两条边从左向右标记为0,1。从树根到树叶的路径上各边标号构成的就是2元前缀码000,001,10,11,01。▌

2. 波兰符号法

定义 设T是一棵根树，对T的每个顶点访问一次且仅一次，称为**行遍T**。

对一棵2叉有序正则树有3种行遍方式：

中序行遍法——左子树，树根，右子树

前序行遍法——树根，左子树，右子树

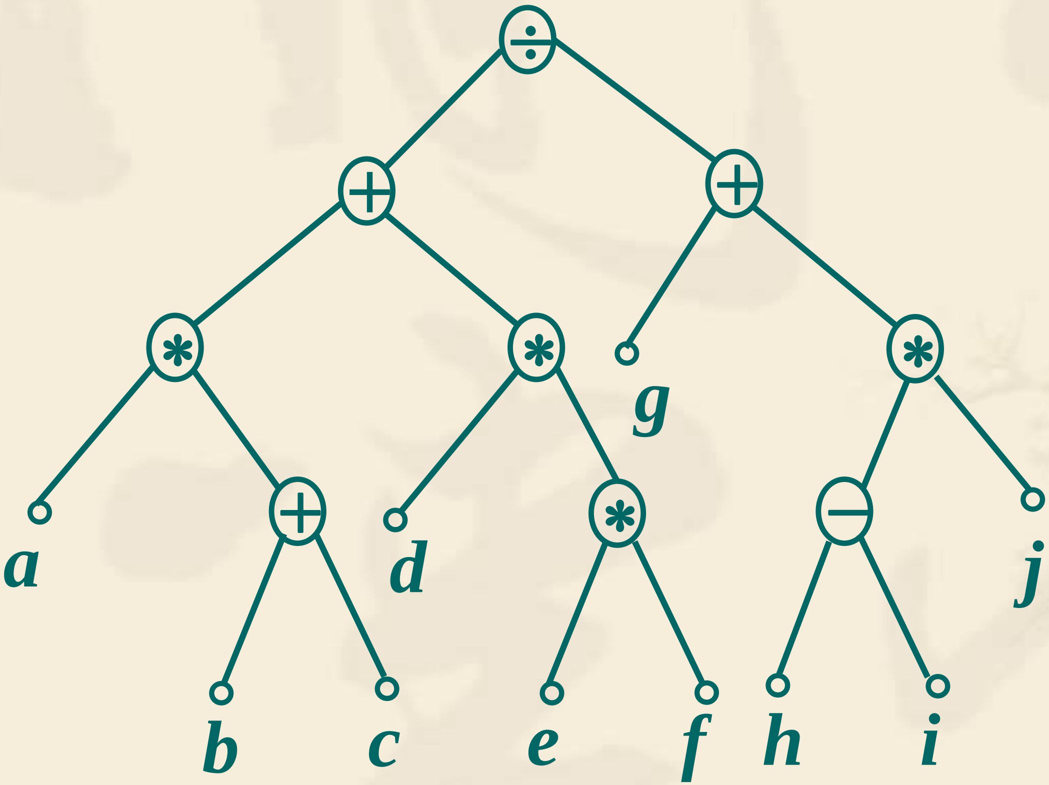
后序行遍法——左子树，右子树，树根

利用2叉有序正则树可以实现二元运算的3种算法。

例 设有算式

$$(a*(b+c)+d*e*f)\div(g+(h-i)*j)$$

用2叉有序正则树表示为



利用前序行遍法，可得算式的另一种表示

$$\div+*a+bc*d*ef+g*-hij$$

计算规则为从右向左，每个算符与其后两个相邻数做运算。

小结：

1. 熟练掌握树的基本概念

   树，森林，树的判别准则

2. 熟练掌握生成树的性质

   生成树，基本回路，基本割集，最小生成树

3. 熟练掌握根树的概念

   根树，2叉树，最优树，前缀码，波兰符号法