

4.2 二叉树的变形 | 4.2.3 哈夫曼树及哈夫曼编码

一、最优树—哈夫曼树

- ◆ 结点的路径长度
- ◆ 树的路径长度

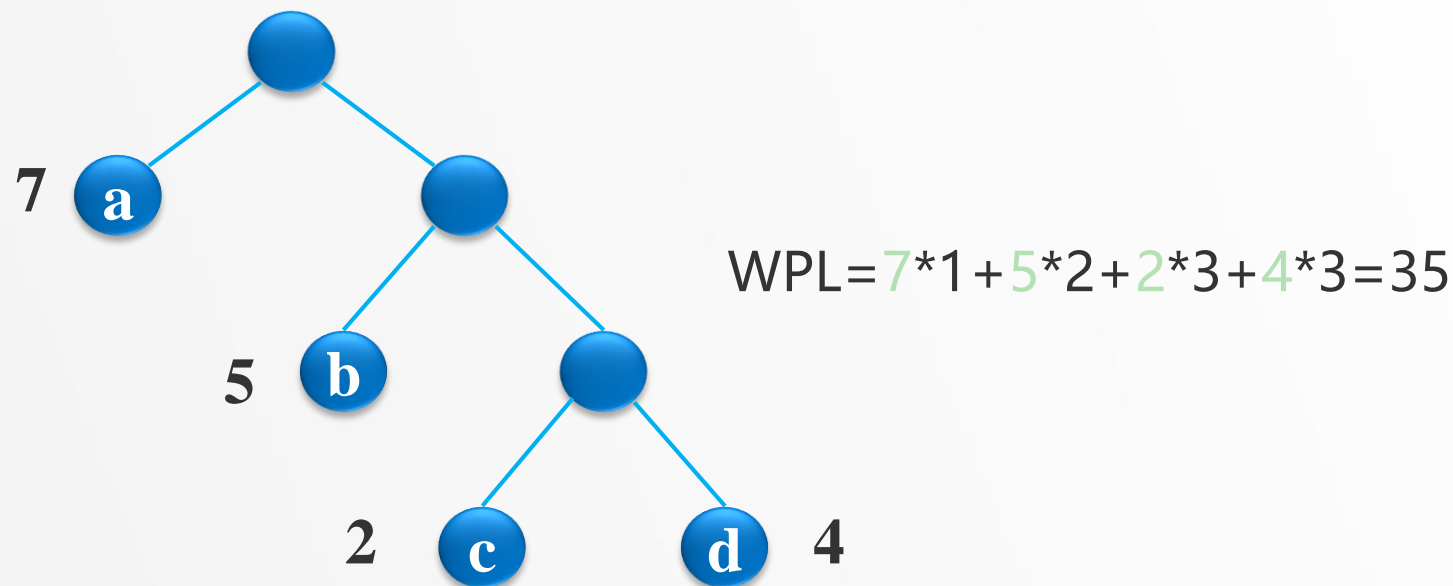
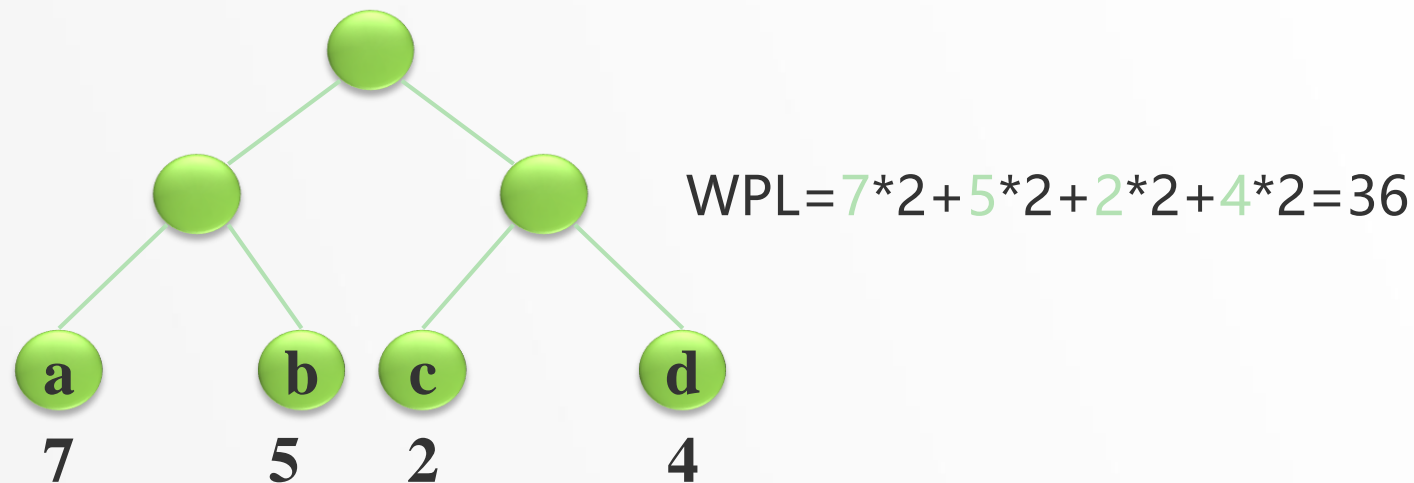
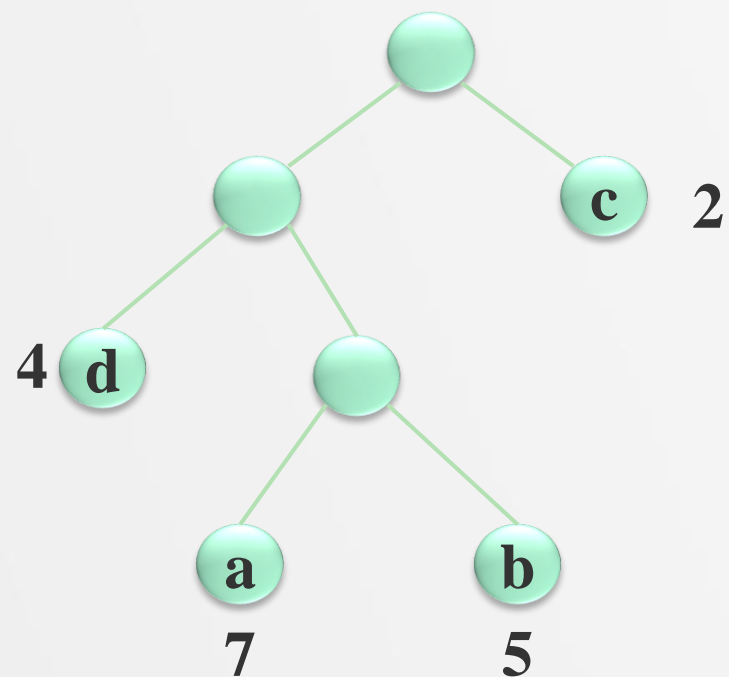
◆ 树的带权路径长度:

$$WPL = \sum_{i=1}^m W_i L_i$$

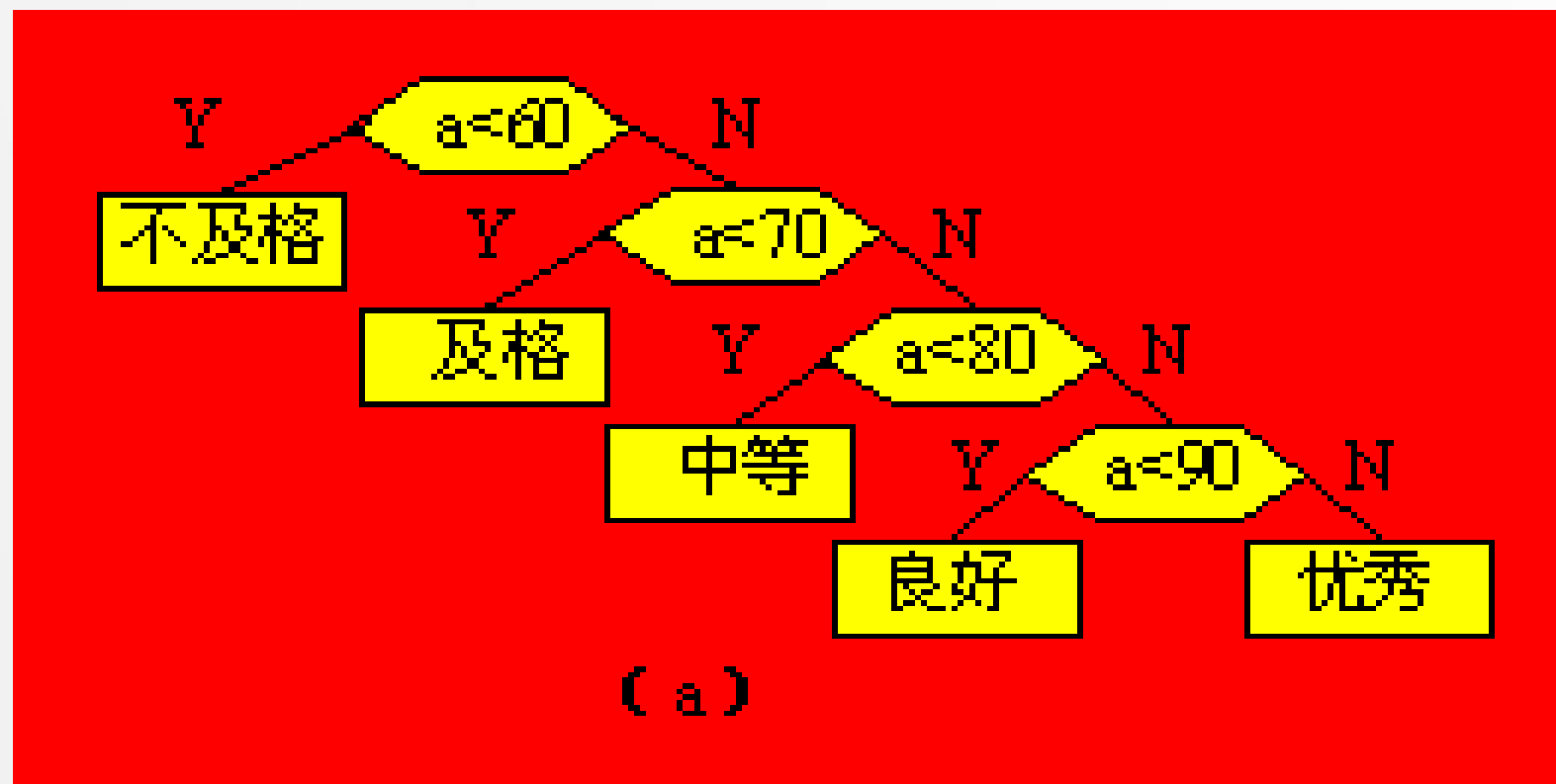
称树的带权路径长度最短的一类树为 “最优树”。

例 有4个结点，权值分别为7，5，2，4，构造有4个叶子结点的二叉树

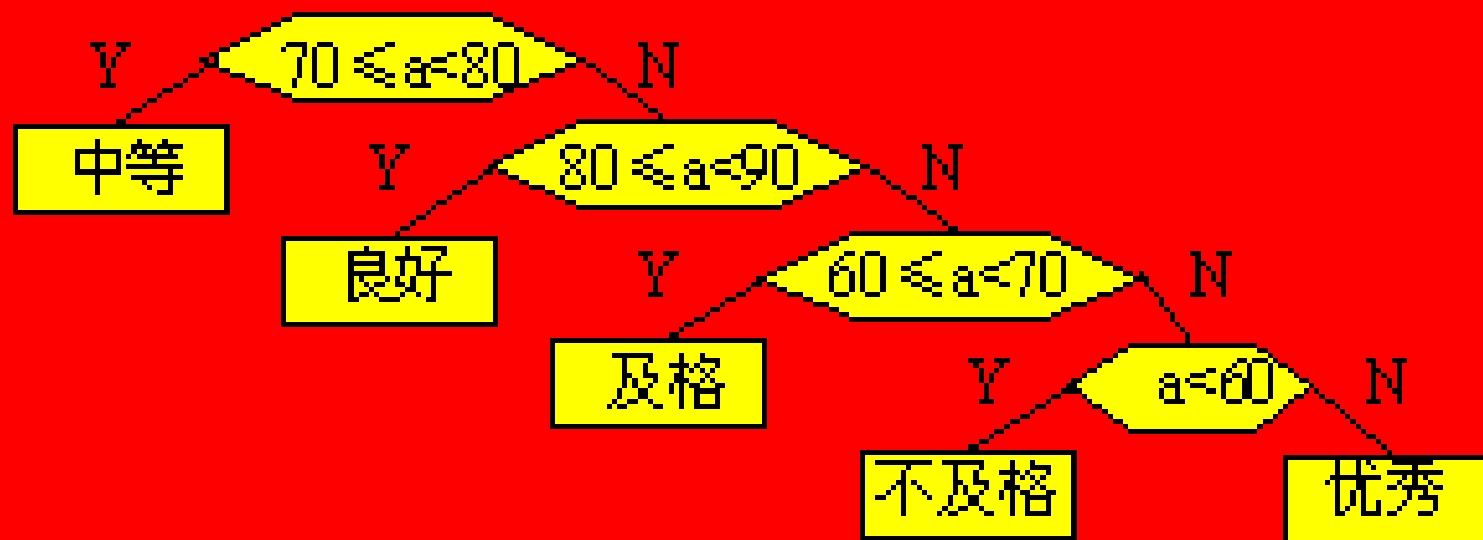
$$WPL = \sum_{k=1}^n W_K L_K$$



在解决某些判定问题时，利用哈夫曼树可以得到最佳的判定算法，例如，要编制一个将百分制转换为五级分制的程序。



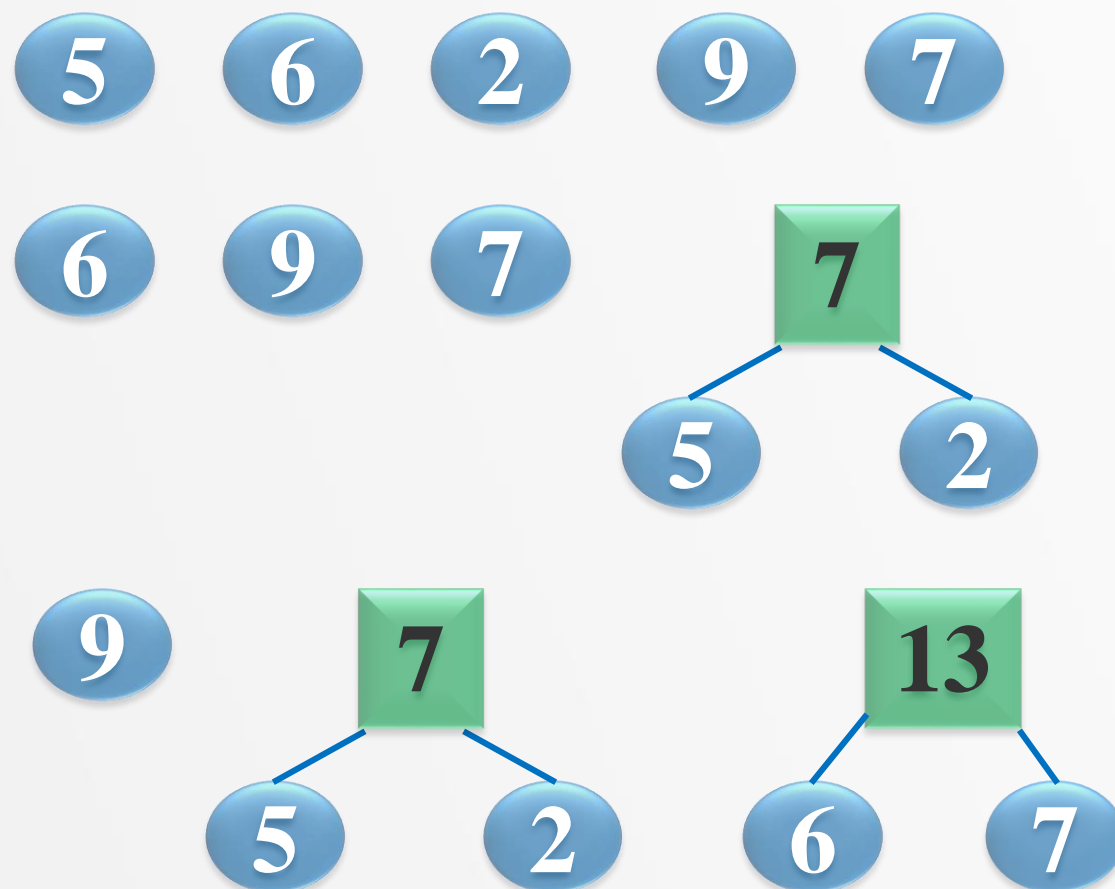
分数	0 - 59	60 - 69	70 - 79	80 - 89	90 - 100
比例数	0.05	0.15	0.40	0.30	0.10

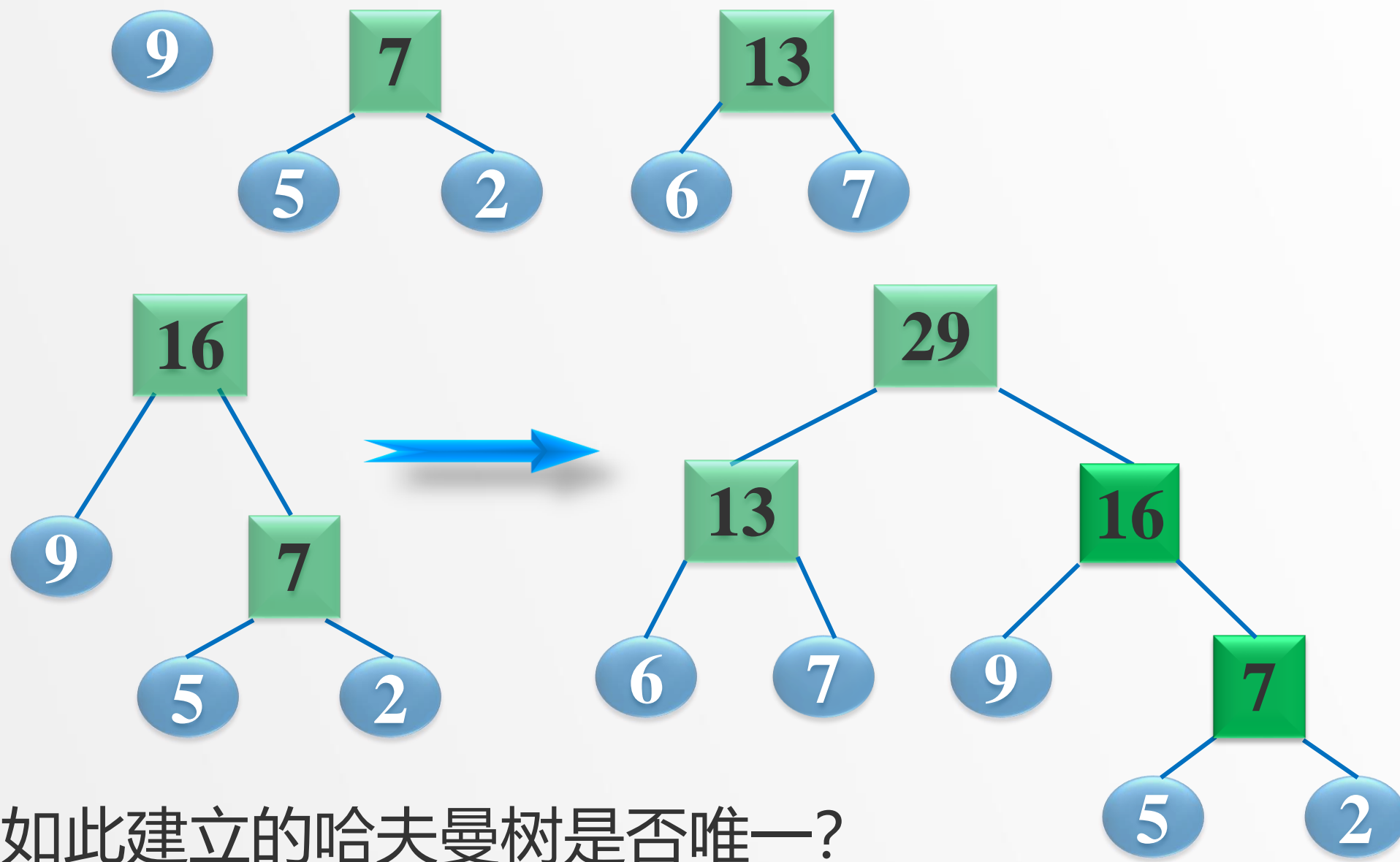


(b)

二、如何构造最优树-- 哈夫曼算法

例如: 已知权值 $W=\{ 5, 6, 2, 9, 7 \}$





讨论

哈夫曼树的任意左右子树是否可以交换？

有 n 个叶子结点的哈夫曼树共有多少个结点？

4.2.3 哈夫曼树及哈夫曼编码

三、哈夫曼编码

例如：需将文字“ABACCD A”转换成电文。
文字中有四种字符，用2位二进制便可分辨。

编码方案1:

A	B	C	D
00	01	10	11

4.2.3 哈夫曼树及哈夫曼编码

三、哈夫曼编码

ABACCD A

编码方案2:

A	B	C	D
0	00	1	01

4.2.3 哈夫曼树及哈夫曼编码

三、哈夫曼编码

ABACCD A

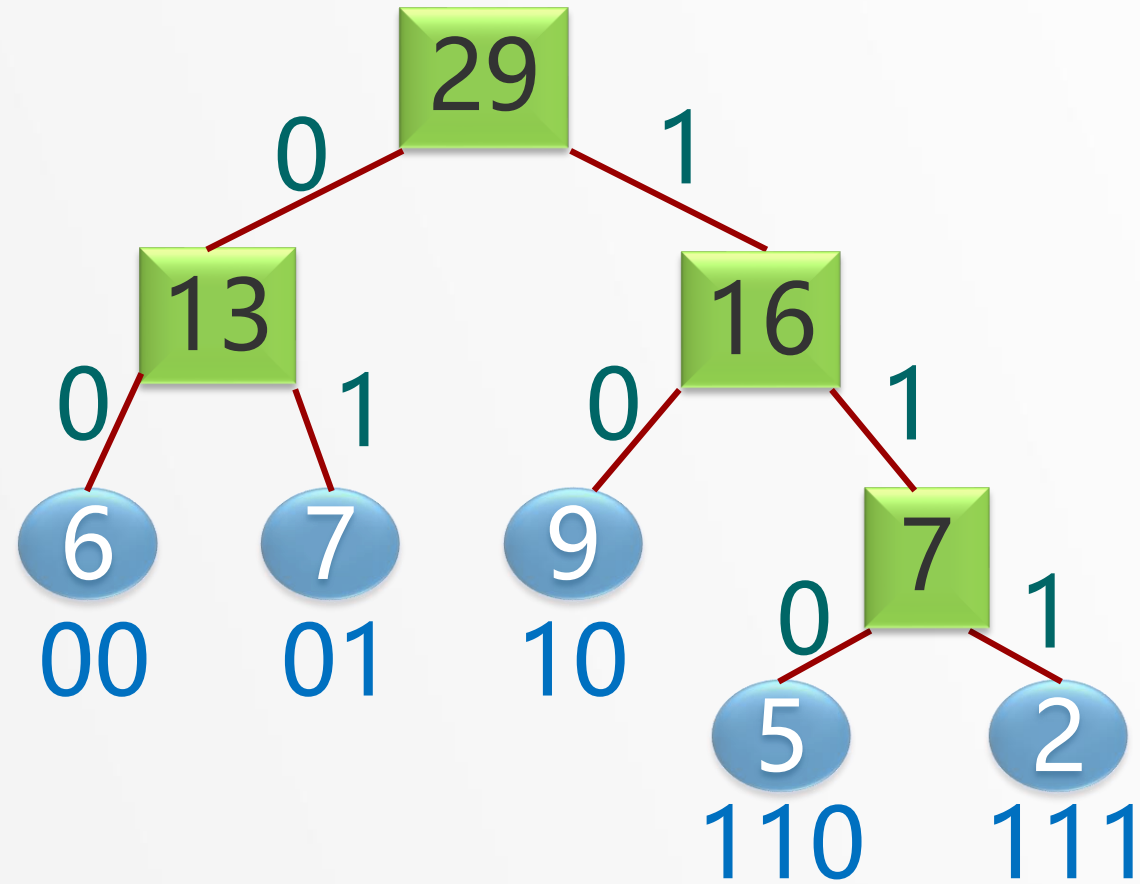
编码方案3:

A	B	C	D
0	110	10	111

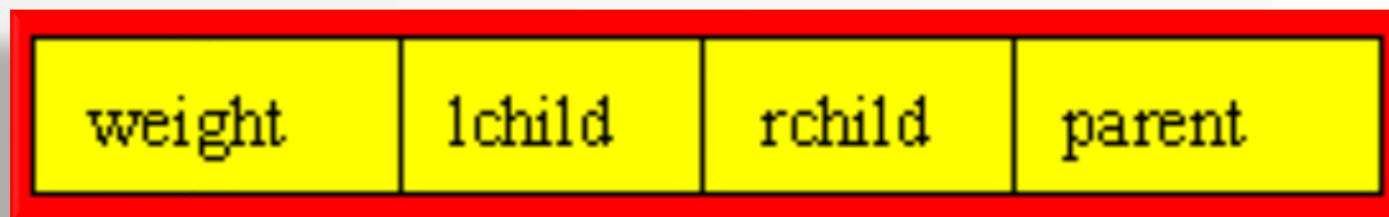
4.2.3 哈夫曼树及哈夫曼编码

三、哈夫曼编码

A:9	10
B:7	01
C:6	00
D:5	110
D:2	111



在构造哈夫曼树时，可以设置一个结构数组HuffNode保存哈夫曼树中各结点的信息，根据二叉树的性质可知，具有 n 个叶子结点的哈夫曼树共有 $2n - 1$ 个结点，所以数组HuffNode的大小设置为 $2n - 1$ ，数组元素的结构形式如下：



哈夫曼编码

作业

给定一组权值 $W=\{8,2,5,3,2,17,4\}$, 画出由此生成的哈夫曼树.

以知一通信系统信源为 $\{a_1,a_2,a_3,a_4,a_5,a_6,a_7,a_8\}$,对应每一信源出现概率为 $\{0.05,0.28,0.07,0.08,0.15,0.23,0.03,0.11\}$,求:

(1) Huffman编码;

(2) 平均码长;