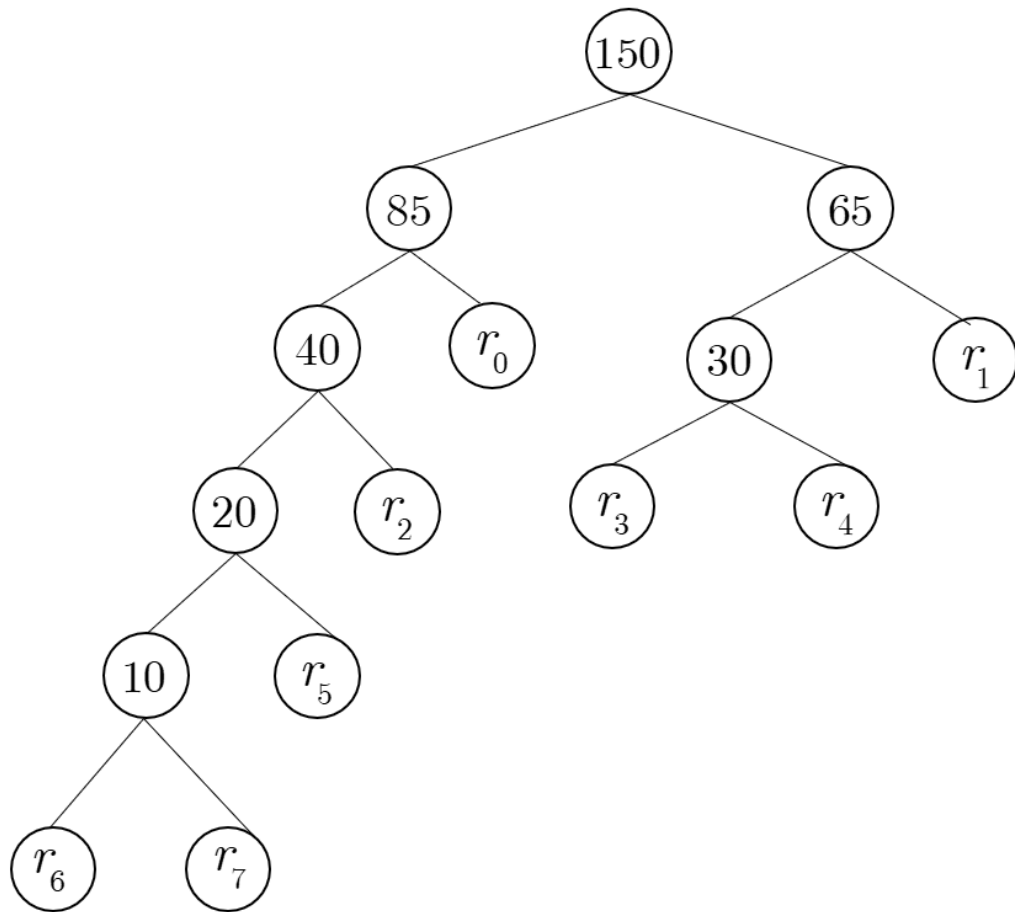


1. 解

(1)

最优二叉树如下图所示



其中对应的编码为

$r_0$  : 01  
 $r_1$  : 11  
 $r_2$  : 001  
 $r_3$  : 100  
 $r_4$  : 101  
 $r_5$  : 0001  
 $r_6$  : 00000  
 $r_7$  : 00001

平均码长为a

$$\begin{aligned}\bar{l} &= \sum_{i=0}^7 p(r_i) l_i \\ &= \frac{45 \times 2 + 35 \times 2 + 20 \times 3 + 15 \times 3 + 15 \times 3 + 10 \times 4 + 5 \times 5 + 5 \times 5}{150} \\ &= \frac{8}{3}\end{aligned}$$

(2)

固定编码码长为  $\bar{l} = 3$ ，那么有

$$R = 1 - \frac{1}{C} = \frac{1}{9} \approx 11.11\%$$

即节约了大约11.11%的存储空间

(3)

### 优势

1. 通过给出现频率较高的符号分配较短的编码，出现频率较低的符号分配较长的编码，该方法能实现较高的压缩率。
2. 是无损压缩算法，得能够完全重建原始数据。
3. 是一种前缀编码，不需要额外的标识符进行标记，解码唯一。

### 局限性

1. 编码较为耗时且计算复杂度高，尤其对于大型数据集的压缩。
2. 需要提前知道不同符号的出现频率，不适合原始数据分布未知或者动态变化的情况。
3. 需要耗费一定空间存储编码表，且查询效率较低，解压缩时需要较长时间搜索和恢复数据。

2. 解

Integer $n$	Parameter $i$	Golomb code $G_{exp}^2(n)$
0	0	000
1	0	001
2	0	010
3	0	011
4	1	10000
5	1	10001
6	1	10010
7	1	10011
8	1	10100
9	1	10101
10	1	10110
11	1	10111
12	2	1100000
13	2	1100001
14	2	1100010
15	2	1100011

### 3. 答

(1)

**信源编码**是为了减少信息冗余度而进行的信源符号变换；

**信源解码**是将信源编码的结果重建为原始的信源符号；

(2)

**信道编码**是为了抵抗信道的噪声和衰减，增加编码冗余度来提高编码的抗干扰和纠错能力；

**信道解码**是根据接收到的信道编码，恢复至原始的信源编码，为最后的信源编码奠定基础；

(3)

**信息量**是事情发生前的不确定性，也是事件发生后能提供的信息量

(4)

**熵**是衡量信源所输出的平均信息量，也是观察信源输出所获得的平均信息量

(5)

**条件熵**是给定一个随机变量的情况下，另一个随机变量的不确定性。

(6)

**互信息**是给定一个随机变量的情况下，另一个随机变量不确定性减少的量

(7)

**信道容量**是所有信源分布中，互信息的最大值。定义了能通过信道可靠传送的最大传输率

海明码 (7, 4) 是一种经典的汉明码 (Hamming Code)，它是一种错误检测和纠正编码方案。

## 编码方法：

### 1. 数据位分配：

- 将要传输的 4 位数据 (D1, D2, D3, D4) 分配到编码字的特定位置上。

### 2. 冗余校验位计算：

- 在一个 (7, 4) 海明码中，有 3 位冗余校验位，将其分别放置在编码字中的特定位置 (P1, P2, P4)。
- 这些校验位的位置是基于奇偶校验，通过位运算计算出来，比如：
  - P1 (位于第1位)：校验数据位D1, D2, D4
  - P2 (位于第2位)：校验数据位D1, D3, D4
  - P4 (位于第4位)：校验数据位D2, D3, D4

### 3. 填充校验位：

- 将计算得到的校验位放置到相应的位置上，形成一个7位的海明码编码字。

## 解码方法：

### 1. 接收数据：

- 接收到一个7位的海明码编码字。

### 2. 计算校验位：

- 对收到的编码字进行校验，通过奇偶校验检查各个位置的数据位和校验位是否符合预期。
- 例如，重新计算 P1、P2、P4 并检查是否存在错误。

### 3. 纠错与恢复：

- 如果检测到错误，可以通过校验位的异或操作来定位并纠正错误的位。
- 例如，如果 P1 出现错误，则知道与 P1 相关联的位是 D1, D2, D4，通过重新计算这三位的校验和来进行纠正。

### 4. 恢复原始数据：

- 从纠正后的编码字中提取出原始的4位数据，即可得到发送方发送的原始数据。