

## 实验三、二叉树的应用 评分细则

总体要求：

- 禁止使用 STL 库，本次试验允许使用 string 头文件
- 限定必须使用 C/C++
- 要求使用二叉树的相关操作完成实验
- 要求源程序可读性好（缩进式，加注释）
- 要求程序健壮
- 要求程序有着较高的时空效率
- 最后的登记打分只进行一次，登记给分之后将不得再次修改（允许同学多次检查，但登记分只进行一次）

评分细则：

- 第一题 6 分，第二题每种文件类型 1 分，实验全部正确完成得 12 分

## 实验内容：

### 题目 1：哈夫曼编码和解码

1. 输入电文字符串
2. 统计电文字符集和每种字符在电文中出现的次数
3. 构建 huffman 树
4. 产生每种字符的 huffman 编码
5. 将电文串翻译成比特流
6. 对电文比特流进行解码

## 题目 2：用 huffman 压缩技术实现对任意文件的压缩和解压缩处理

要求对所有的文件类型进行压缩，压缩之后的文件后缀名为 huff。同时，可以对所有后缀名为 huff 的压缩文件进行解压缩。

### 题目具体要求：

因为实际上题目 2 包含题目 1，所以如果完成题目 2 则直接默认题目 1 正确，题目 2 详细要求如下

1. 题目 2 中以 1 个字节(8bit)为单位进行 huffman 编码
2. 对任意文件进行压缩后可以输出一个后缀名为 huff 的单文件，并且可以对任意一个后缀名为 huff 的单文件进行解压还原出原文件。
3. 群内将提供 5 不同种类的文件包括文档、图片、视频、可执行文件等进行压缩测试，要求可以完成压缩和解压的步骤，并且解压出来的文件没有任何损失。

### 如果未完成题目 2，题目 1 详细要求如下

1. 题目 1 中以 1 个字符为单位进行 huffman 编码
2. 编码时要求输出如下信息

原始字符	编码后字符
aaabbbbbbbccccddddd.....	一串乱码.....
频率信息	编码信息
a 5000 b 20000 c 8000 d 4000 e 27000 .....	a 1100 b 111 c 1101 d 0 e 10 .....

3. 解码时要求通过输入“编码后字符”和“编码信息”，完成解码输出原始字符。
4. 群内将提供 2 个文本文件进行测试。