

# Huffman System

---

吕艺

517021910745

November 25, 2018

## 1 需求分析

1. 本演示文件中操作的主要数据结构为哈夫曼树，数的叶子个数不限。首先需要在主菜单中输入命令'I'来进行哈夫曼树的初始化(这是之后所有操作的前提条件)为增强程序的健壮性，用户如果想在哈夫曼树初始化前就进行其他操作显示错误信息并被要求重新输入。初始化操作的输入为字符个数 $n$ (输入结束后按回车)，字符(不带空格输入，输入结束后按回车)和他们对应的权重(每个整数之间以空格分割，输入结束后按回车键)之后可进行的操作有编码('E', 根据已建的哈夫曼树对给定文件中的字符)，译码('D', 利用已建的哈夫曼树将给定文件中的二进制码还原为字符)，打印代码文件('P', 打印在屏幕上并保存到文件中)和打印哈夫曼树('T', 打印在屏幕上并以凹入表的形式保存在文件中)。

2. 演示文件以用户和计算机的对话方式进行，即在计算机终端上显示合适的提示信息之后，由用户在键盘上输入演示程序中规定的运算命令；命令结束完后，生成相应的文件或屏幕上显示结果。

### 3. 程序执行的命令包括:

1. I: 初始化 (Initialization)。从终端读入字符集大小  $n$ ，以及  $n$  个字符和  $n$  个权值，建立哈夫曼树，并将它存于文件 `hfmTree` 中。
2. E: 编码 (Encoding)。利用已建好的哈夫曼树，对文件 `plainFile` 中的正文进行编码，然后将结果存入文件 `codeFile` 中。
3. D: 译码 (Decoding)。利用已建好的哈夫曼树，对文件 `codeFile` 中的代码进行译码，然后将结果存入文件 `textFile` 中。
4. P: 打印代码文件 (code Printing)。将文件 `codeFile` 显示在终端上，每行 50 个代码。同时将此字符形式的编码文件写入文件 `codePrint` 中。
5. T: 打印哈夫曼树 (Tree printing)。将哈夫曼树以直观的方式 (树或凹入表形式) 显示在终端中，同时将此字符形式的哈夫曼树写入文件 `treePrint` 中。

### 4. 测试数据

1. 已知某系统在通信联络中可能出现 8 种字符，其出现频率分别为 0.05, 0.29, 0.07, 0.08, 0.14, 0.23, 0.03 和 0.11。
2. 用下表给出的字符集和频度的实际统计数据建立哈夫曼树，并实现以下报文中的编码和译码: "THIS PROGRAM IS MY FAVORITE"。

Table 1.1: 测试数据

char	space	A	B	C	D	E	F	G	H	I	J	K	L	M
Weight	186	64	13	22	32	103	21	17	47	57	1	5	33	20
char	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	
Weight	57	63	15	1	49	51	80	23	8	19	1	16	1	