

实验三 用赫夫曼树对文件进行编码和解码

一、描述

编写一个程序，实现如下功能：

- 1) 编码功能：输入一个文件 A，依据 A 文件的内容，对文件中的每个字节采用赫夫曼算法进行重新编码，并将编码所用的赫夫曼树和对文件内容重新编码的结果输出到另外一个文件 B 中。
- 2) 解码功能：输入编码后的文件 B，解码还原，将还原内容保存在文件 C 中，要求文件 A 和 C 内容完全一致。

二、题目要求：

- 1、输入文件 A 可以是任意文件，即支持文本文件和二进制文件。
- 2、输出文件 B，需包含用于支持将文件 B 的内容还原为文件 C 所需的必要信息。
- 3、编码时，可以将一个字节，编码为字符'0'和'1'组成的字符串；输出到文件 B 时，可以输出这个字符串来替代 01 的比特串。
- 4、解码时，读入 B 中的信息还原赫夫曼树并对内容进行解码后输出到文件 C。
- 5、选做，要求输出的文件 B 中的编码结果用 0/1 的 bit 进行输出。
- 6、选做，可以和其他小组协商 B 文件的格式，在共同的文件格式约定下，支持到小组间程序间的互操作：即可以用小组 1 的程序进行编码，用小组 2 的程序进行解码，反之也可以。
- 7、程序实现要求：
 - a) 可执行程序方面：
 - i. 可以实现为两个程序，即分为单独编码程序和单独解码程序；
 - ii. 可以实现为一个程序，程序在一次运行时，完成编码或解码的一个功能，具体完成的功能可以由命令行参数或输入参数控制；
 - b) 输入：
 - i. 输入必要的参数，包括输入文件名称、输出文件名称和可选参数
 - ii. 可以用命令行参数方式实现，也可以用提示后输入方式实现。
 - c) 输出：
 - i. 编码：输出编码后的文件 B
 - ii. 解码：输出解码后的文件 C
 - d) 程序实现要求：
 - i. 自行实现赫夫曼树相关数据结构和算法；
 - ii. 文件读写操作使用标准库或其他库和类库；
 - iii. 如需进行排序，可使用标准库的排序算法或自行实现排序；
 - iv. 用 C++ 的类来实现的同学，要练习如何使用参数在调用者和被调用者之间进行数据传递。
 - e) 鼓励采用多文件方式编写程序。
- 8、给出测试用例的设计说明
 - a) 至少要测试 1 个文本文件和 1 个二进制文件。
 - b) 说明每个测试用例的输入、预期输出
 - c) 说明测试用例是否能覆盖所有测试场景
 - d) 要求考虑如何用最少的测试用例来完成测试。

- 9、给出数据结构设计说明，至少包括：
- a) 如采用教材中的赫夫曼相关的数据结构，无需再给出设计说明
 - b) 如果在教材中的赫夫曼数据结构基础上有增加和改变，需给出相关说明；
 - c) 给出文件 B 的结构设计说明，包括：存储了那些必要信息，存储格式。
- 10、 给出程序说明，用文字描述程序代码主要处理流程和功能，赫夫曼相关算法，如对教材算法无大的改变，可以不用做说明。
- 11、 分工要求：本次实验建议两位同学编写程序，一位同学编写报告和进行测试。
- a) 两位同学的分工为：1 位同学完成编码相关功能；1 位同学完成解码相关功能。
 - b) 讨论要求：现有实现的那些方面可以进行改进来提高程序编码和解码的速度，将讨论结果写入实验报告。