

Project--压缩解压工具

背景介绍

文件压缩是节省磁盘存储和帮助减少传输时间的一个重要途径。

一般来说，有两种类型的压缩机制——有损压缩和无损压缩。顾名思义，有损压缩在恢复压缩文件后会导致内容丢失，通常用于媒体数据，而无损压缩可以完全恢复原始文件。计算机中常见的压缩格式，如zip、rar等，都属于无损压缩方式。

在这个Project中，你需要使用**Huffman编码**来实现压缩和解压工具。

提交后，还需要进行**线下面试**。面试时间另行通知。

要求：

- 语言：尽量使用C++。如有特殊情况，请单独联系TA说明。
- 自定义输入、输出方式。要求线下面试时能清楚、流畅、便捷、快速地压缩给定的Test cases文件夹或其**任意指定子集（包括单个文件、文件夹）**，并能完整解压出源文件。
- 提交文件：
 - **项目源代码**
 - **开发文档**
 - 项目文件结构说明
 - 项目设计、实现的具体思路
 - 性能测试结果（如每个测试用例的压缩率、压缩解压时间）
 - 遇到的问题和解决方案（项目重难点）
 - 优化技巧（项目亮点）
- **Bonus**：比较你的解决方案和其他压缩工具（如WinRAR，7Zip，HaoZip等）的压缩时间和速率，思考差异的原因

提交方式和时间

- 提交方式：请将所有文件打包压缩，并命名为“**学号-姓名.zip**”，上传至elearning对应入口处。
- **Final DDL：2021年12月10日23: 59**
- **中期检查**：为了防止赶DDL的情况，拟定于**11月22日lab课**进行中期检查，需要所有同学出席，一对一地向TA介绍大概的实现思路和完成情况。希望所有同学认真对待，提前规划好项目时间安排。

评分标准

内容	分数
非空文件——压缩&解压	20
非空文件夹——压缩解压	20
空文件——压缩&解压	10
空文件夹——压缩&解压	10
性能（时间&效率）	15
自定义输入输出便捷度、用户友好程度	10
开发文档	10
优化 & bonus	5

一些建议

- 仔细思考
 - 如何存储霍夫曼树，以便在解压时有效地恢复
 - 如何恢复字符，因为他们的编码长度不是固定的
- 许多文件类型已经被压缩，比如JPEG和MPEG。可以尝试使用标准工具对这些文件进行压缩，看看会发生什么。
- 运行时间与计算机硬件有一定的相关性，因此请记住在进行效率测试时要包括计算机的配置。
- 避免串接或字符串减法可以帮助减少运行时间。
- 在压缩和减压时要注意EOF。
- 注意空文件和空文件夹。
- 灵活使用调试器，记得写注释。
- 如果遇到任何问题，首先尝试自己思考，然后在网上寻找解决方案。
- 如果对本Project有任何不清楚的地方，请与TA联系。
- **禁止抄袭任何人（不限于班级同学）的代码。一经发现，抄袭者和被抄袭者皆按0分处理。**

- 性能测试结果可用excel表格记录，示例如下图：

	A	B	C	D	E	F	G
1	test case	Compression time/ms	Compression Size/bytes	Origin Size/bytes	Other tool Size/bytes	Compression Rate	Decompression time/ms
2	1.1						
3	1.2						
4	1.3						
5	1.4						
6	1.5						
7	1.6						
8	1.7						
9	1.8						
10	1.9						
11	1.10						
12	1.11						
13	1.12						
14	1.13						
15	1.14						
16	1.15						
17	1.16						
18	1.17						
19	1.18						
20	1.19						
21	1.20						
22	1.21						
23	1.22						
24	1.23						
25	1.24						
26	1.25						
27	1.26						
28	1.27						
29	1.28						
30	1.29						