Project 1

File Compression Using Huffman Coding

Zheran Fang, 13212010002@fudan.edu.cn

Oct 8, 2014

General Information

- Implement a file compression tool using Huffman Coding
- In Java
- Handout: Oct. 8, 2014
- Available on FTP `PROJECT` directory

Introduction

- In computer systems, information is encoded in bits.
- We need data compression
  - Save storage space
  - Simplify file transfer
- Data compression
  - Lossy compression
  - Lossless compression (this project)

File in Bits

- View file as a sequence of bits
- Each character is encoded into bits
  - ASCII code (7 bits)
  - GB2312 (16 bits)
  - N-bit coding (2^n bits)
- Use Huffman Coding to reduce bit length

How to Compress?

- Read file and count character
- Build Huffman Tree (optimal prefix code)
- Encode the file using Huffman Coding
- Save into file (including the tree)
- Decompress?

Requirement

- Single file compression & decompression
- Folder compression & decompression (recursive and non-recursive)
- User-friendly GUI and User Manual
- Development Document (including test cases)
- *Optional*: Optimize your tool and compare with WinZip, 7-zip, bzip, etc.

Grading

Part	Percentage
Single file compression & decompression	30%
Folder compression & decompression	30%
UI design & User Manual	15%
Performance (Time & Rate)	10%
Development Documentation	15%
Optimization & Comparison	10% *(bonus)*

Submission

- Deadline: Oct. 28 2014, 23:59 (GMT+08:00)
- Package your project
  - Include an executable jar file
  - Using standard compression format, not yours :)
- Submit to FTP
- Face-to-face interview

Policy

- No cheating
- No late policy

Thanks!

Q&A

Feel free to contact me via E-mail or WeChat