

Cosc 2P03 Fall 2014

Bonus Assignment

(Due at start of 2P03 Exam)

Bonus Assignment

This is a bonus assignment and is optional. If completed, it will be worth 5% of your grade, and replace the lowest percentage mark of the assignments in this course.

LZW compression

For this assignment you will be implementing the LZW compressor and de-compressor. Let's not reinvent winzip but concentrate on the algorithm. Consider a [data file](#) composed of 0's, 1's, blanks (char 32) and the EOL character. Hence the alphabet is 4 characters in length.

Part A

File compare utility

To be sure the compressor and de-compressor work write a simple file compare utility. This utility will prompt for 2 files and compare them based on the above alphabet. Thus only characters of the alphabet need to be considered, all others can be ignored. If the 2 files are the same, then issue a message that the files are the same, otherwise issue a message that they are different.

Part B

Write a compressor which will compress the data file to a list of integers which represent the encoded file. To aid in the look up of character strings implement a hash table using external chaining. The size of the table will be 301. Consider a code string **C** composed of characters C_i sequentially numbered where $i \geq 1$. Thus the following hash function may be used.

$$f(C) = \text{SumOf}(87 * i * \text{ord}(C_i)) \bmod \text{TableSize}.$$

The decompressor will read the integer file and recreate the original data file. Use the file comparator from Part A to verify things are working.

Output

In addition to the file comparison utility, print a copy of the code table, after compression and after decompression.

Submission Requirements:

- **Cover Sheet** completely filled out, available from: "<http://www.cosc.brocku.ca/coverpage>" Note: your assignment will not be marked unless one is submitted with the assignment on the assignment due date.
- **Commented and properly documented** source code listing.
- Listing of your input and output.
- Source code is to be Java.
- Electronic submission, run the script "submit2p03" from sandcastle.
- Statement on coversheet with following information.
 - Platform, e.g. Mac, PC, Commodor 64, my Java enabled wrist watch.
 - Compiler Version, e.g. Java 1.6, Java 1.7 e.g.

Good Luck