|  |  |  |  |
| --- | --- | --- | --- |
| **Instructor** |  | **Due Date** |  |

**PROJECT Data Compression: Huffman Coding**

**Objective** To investigate the mechanisms of data compression of text files.

***PROJECT DESCRIPTION***

Read and review the " Focus on Data Compression " details regarding Huffman Coding in Chapter 7 of your course textbook.

Implement the Huffman Coding algorithm on the following passage by using a hand trace of the Huffman Coding tree or a suitable software tool such as MS Paint, MS Excel, MS PowerPoint, Adobe Photoshop CC , etc.

SUPERCALIFRAGILISTICEXPIALIDOCIOUS

• Find the number of bits required to encode the message.

• Compute the lower bound of compression based on the following formula:

*n*

*E* = ∑ { − *P* ( *x* *i* ) × log [ *P* ( *x* *i* ) ] }

*i* = 1

What is the redundancy ( in bits ) given this computation? Refer to the sample computational sheet that follow.

***Information about This Project***

The Huffman Coding routine examines each symbol and its frequency of occurrence and encodes each symbol using the frequency of each symbol such that those with higher occurrence ( i.e. weights ) have fewer bits in their encoding.

**[ Web Link for Further Information ]**

[**https://www.geeksforgeeks.org/practice-questions-on-huffman-encoding/**](https://www.geeksforgeeks.org/practice-questions-on-huffman-encoding/)

***Steps to Complete This Project***

**STEP 1 View the Video Lecture**

After you review the course textbook on the topic of data compression, view a video on this subject such as that given in the Web link below.

**[ How Computers Compress Text: Huffman Coding and Huffman Trees ]**

[**https://www.youtube.com/watch?v=JsTptu56GM8**](https://www.youtube.com/watch?v=JsTptu56GM8) **TRT 6:30**

**STEP 2 Launch a Drawing Tool or Create the Tree by Hand**

Draw the Huffman Tree by hand and take a photo camera snapshot of the drawing or use suitable digital software to create the tree.

**STEP 3 Perform the Tasks of the Project**

After you design the tree, perform any other tasks required by the project and answer any other questions that appear.

**STEP 4 Submit the Project**

Submit a copy of your completed project for credit.

**STEP 5 Questions and Answers Concerning this Computer Laboratory Project**

Open MS Word and, within a new document, place your responses to these questions. Submit your completed MS Word document for credit.

**PROJECT Data Compression: Huffman Coding**

**(1) ( Huffman Coding )**

What are some applications of the Huffman Coding Process?

|  |
| --- |
|  |

**(2) ( Greedy Methods )**

Visit the Web link below to review the given article and then respond to the following question.

**https://www2.cs.duke.edu/csed/poop/huff/info/**

Explain why the Huffman Coding Method is considered as a " greedy " process?

|  |
| --- |
|  |

**(3) ( Lossy versus Lossless Compression )**

Visit the Web link below to review the given article and then respond to the following question.

[**http://mathworld.wolfram.com/HuffmanCoding.html**](http://mathworld.wolfram.com/HuffmanCoding.html)

Is Huffman Coding considered a lossy or lossless compression process?

|  |
| --- |
|  |

**(4) ( Binary Trees )**

Is the Huffman Coding Method also a type of Binary Tree, from data structures? Explain your answer.

|  |
| --- |
|  |

**(5) (** **Variable - Length Encoding )**

Explain how Huffman Coding is related to variable - length encoding.

|  |
| --- |
|  |

**PROJECT Data Compression: Huffman Coding**

**STEP 6 The Computation Worksheet for the Lower Bound of Compression**

Here is a sample worksheet with one letter that will assist you in computing the lower bound of compression of this process.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Letter | Frequency | *P* ( *x* *i* ) | log [ *P* ( *x* *i* ) ] | − *P* ( *x* *i* ) × log [ *P* ( *x* *i* ) ] |
|  |  |  |  |  |
| D | 1 | 1 / 34 | −1.5315 | 0.0450 |
|  |  |  |  |  |
|  |  |  |  |  |