|  |
| --- |
| **“DJ” IT PROJECT** |
| Huffman code, Hamming code |
|  |
| Team: Daniyar Kuntubayev Zhahanger Khusainov |
|  |
| **Astana IT University** |
| **06.06.2021** |
|  |

Outline:

* Introduction
* Algorithms:
  + Huffman code
  + Hamming code
  + Entropy
* GUI application
* Reflection
* Conclusion
* References

Introduction

Nowadays, there are many types of data transmission and also different types of data, large and small, categorical and numerical, etc. But in all, in order to correctly transfer, different compression algorithms are used. Such as Shannon-Fano, Huffman, Hamming, etc. Maybe the reader will have a question here? What are the algorithms? How do they work? What are they needed for? Here in today's project we will try to answer your questions by analyzing the algorithm of Huffman and others.

Algorithms

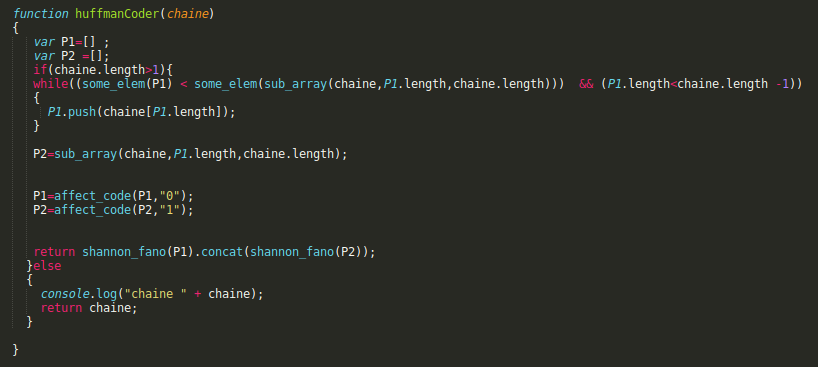
1. Huffman code

Firstly, little bit about algorithm. Huffman code is a specific type of optimal prefix code that is used for lossless data compression. The process of finding or using such a code proceeds by means of Huffman coding, that means an algorithm developed by David A. Huffman. He published them in the 1952 article "A Method for the Construction of Minimum-Redundancy Codes" when he was a Sc.D. student at MIТ.

Secondly, lets talk about how to work this algorithm. Huffman coding assigns codes to symbols such that the length of the code depends on the relative frequency or weight of the corresponding symbol. Huffman codes are variable in length and do not contain prefixes (no code is a prefix of any other). Any binary code without prefixes can be represented as a binary tree with encoded characters stored on leaves.

Let's look at Huffman Code in two cases: Encoding and Decoding.

Encoding:



Decoding:

To decode an encoded string, follow the zeros and ones to the sheet and return the character there.You are provided with a pointer to the root of the Huffman tree and a binary string to decode. You need to print the decoded string.



1. Hamming code

Firstly, little bit about algorithm. Hamming codes are a family of linear error-correcting codes. Hamming codes can discover up to two-bit errors or correct one-bit errors without finding uncorrected errors. The simple parity code cannot correct errors and can detect only an odd number of bits in error. Hamming codes are ideal codes, that is, they achieve the highest possible rate for codes with their block length and minimum distance of three.  
Richard W. Hamming invented Hamming codes in 1950. That looks like a way of automatically correcting errors introduced by punched card readers. In his genuine article Hamming develop his general idea, but focused on the Hamming(7,4) code which complements three parity bits to four bits of data.

Secondly, lets talk about how to work this algorithm.

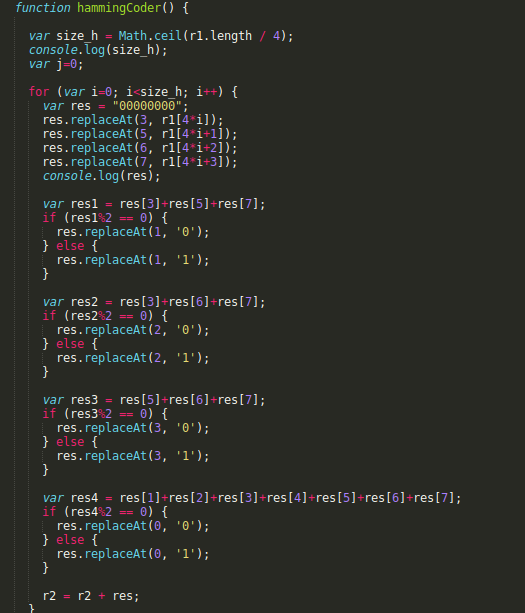
Encode:

The process of encoding a message using Hamming code is 3 steps.

Step 1. First, we calculate the number of redundant bits.

Step 2: Place them in exact / correct position.

Step 3: Let's calculate their values.



Decode:

The Hamming decoding process includes the following steps. This process is a recalculation to detect and correct errors in the message.

Step 1: Count the number of redundant bits

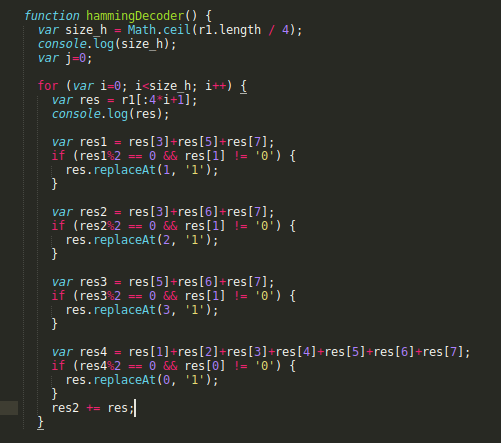
Using this formula, we encode the redundant bits:

2p ≥ n + p + 1

Step 2: fix the position of all redundant bits

redundant bits in the number of 'p' are placed in the bit positions of the power of 2, for example, 1,2,4,8,16,32, etc.

Step 3: parity check (odd parity and even parity)

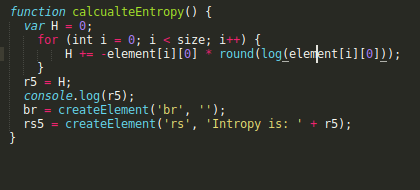


1. Entropy

Firstly, little bit about algorithm. Entropy coding refers to an encoding in which the entropy of the compressed data coincides with the entropy of the original source, and the compressed data can be used to completely restore the original information.  
  
One of the main types of entropy coding creates and assigns a unique code without prefixes to each unique character that appears in the input. These entropy encoders then compress the data by replacing each fixed-length input character with a corresponding unprefixed variable-length output codeword. The shortest codes are used for the most common characters, since the length of each codeword is approximately proportional to the negative logarithm of the probability of that codeword occurring.

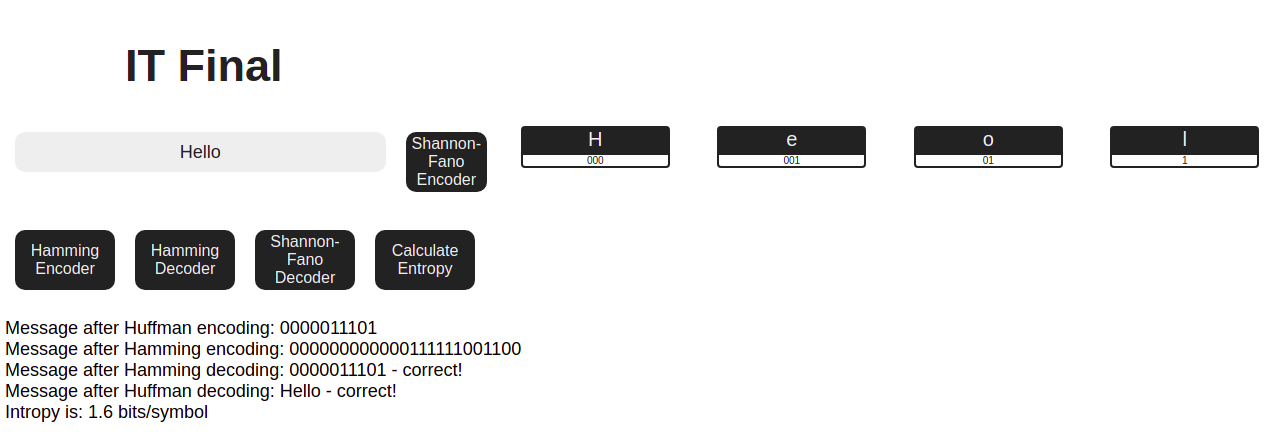
Secondly, lets talk about how to work this algorithm.

He goes through the element, and calculates their ratio as binary.



1. GUI application

Our application was written in the Javascript language. It has a form input and also buttons with different functions. You can write which word or sentence and select the algorithm you need such as Huffman encoding and decoding or entropy. After fetching the algorithms, it will display the results in a list.



1. Reflection

Daniyar Kuntubayev: Huffman code, GUI APP

Zhahanger Khusainov: Hamming code, Entropy

Presentation:

<https://www.canva.com/design/DAEgnSHFj5s/share/preview?token=P9ZRVlv5TId1RX2chj0xIw&role=EDITOR&utm_content=DAEgnSHFj5s&utm_campaign=designshare&utm_medium=link&utm_source=sharebutton>

Conclusion

After all the work done, we think we were able to answer your questions. Compression algorithms play a very important role in our life. In the end, we want to thank Huffman for what he create and Miss Elvira for teaching us.

References:

https://en.wikipedia.org/wiki/Huffman\_coding

https://en.wikipedia.org/wiki/Hamming\_code

<https://en.wikipedia.org/wiki/Entropy_encoding>

https://www.geeksforgeeks.org/hamming-code-in-computer-network/