Huffman Coding

Huffman coding is a lossless data compression algorithm. The idea is to assign variable- length codes to input characters, lengths of the assigned codes are based on the frequencies of corresponding characters.

The most frequent character gets the smallest code and the least frequent character gets the largest code.

The variable-length codes assigned to input characters are Prefix Codes, means the codes (bit sequences) are assigned in such a way that the code assigned to one character is not the prefix of code assigned to any other character.

Huffman coding is used for :

- conventional compression formats like GZIP, BZIP2, PKZIP, etc.
- For text and fax transmissions.
- In statistical coding

There are quite a lot of real-world applications of Huffman Encoding. ZIP is perhaps the most widely used compression tool that uses Huffman Encoding as its basis. The latest of the most efficient lossless compression algorithms, Brotli Compression, released by Google in 2020 also uses Huffman Coding.

Huffman is also used in all the mainstream compression formats, from GZIP, PKZIP (winzip etc) and BZIP2, to image formats such as JPEG and PNG.

Huffman coding is one of the greedy algorithms widely used by programmers all over the world. It is one of the best ways to compress the data which losing it and transfer data over the network efficiently. It is highly recommended to understand the working of Huffman coding and use it to compress your data efficiently.