University College Dublin

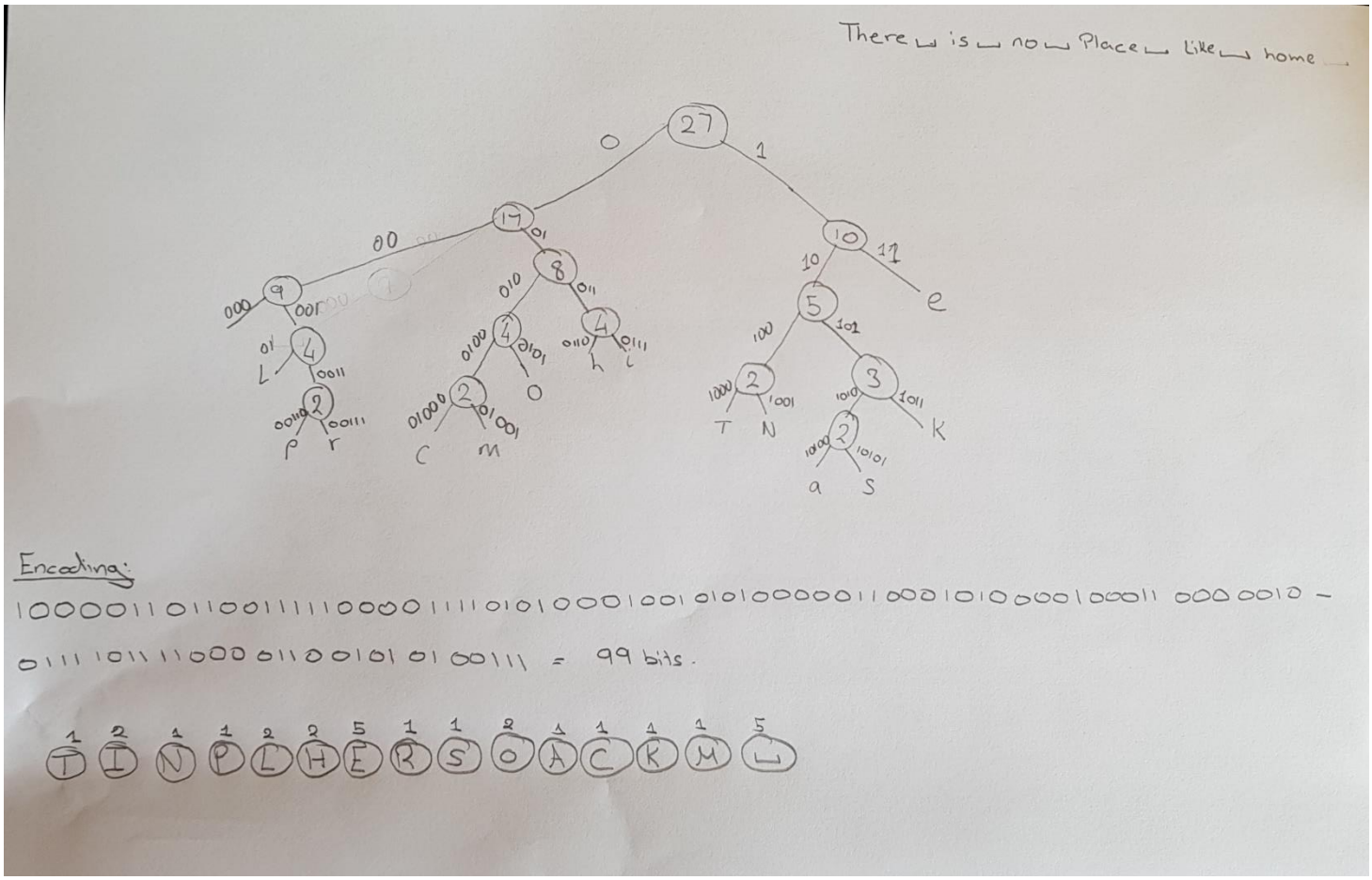
Coding Assignment - COMP20290

Aman Parary - 18417714

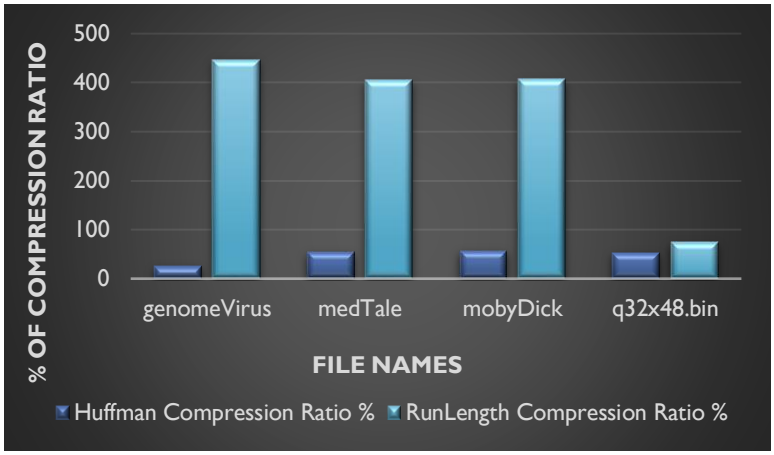
3/5/2020

Codeword Table

Key	value
_	000
A	10100
C	01000
E	11
H	0110
I	0111
K	1011
L	0010
M	01001
N	1001
O	0101
P	00110
R	00111
S	10101
T	1000



Huffman compression is a lossless algorithm. It compresses a file by deleting redundancies in the file. When I compressed a file twice, I saw no changes in the file size (bits). I believe this is because a compressed file does not have any more redundancies, so compressing the file again won't reduce the file size.



The bar chart compares the ratio percentage of the given data files using Huffman and RunLength Algorithm. RunLength compression ratio has a higher percentage when compressing genomeVirus, medTale and mobyDick files. RunLength is not suitable for ASCII text because it needs multiple long runs of 1s and 0s, hence the higher percentage ratio.

The 'q32x48.bin' file % ratio for RunLength is not far off compared to Huffman. The Huffman compression uses 29% fewer bits than RLE. The compression time for RLE takes longer than Huffman by 523402 ns. I believe the reason is because 00000000 and 11111111 characters are more likely to happen in text files, therefore, the characters have a possibility to be encoded with 2 or 3 bits which results to a significant compression.

Bitmap test case (1536 bits 4365101 ns)

Compressed

```
$ java helper_code/HuffmanAlgorithm - < data/q32x48.bin > data/q32x48.bin.huf1
$ java helper_code/BinaryDump 0 < data/q32x48.bin.huf1
816 bits                compression ratio 816/1536 = 53%
Time taken: 2235399 ns.
```

Decompressed

```
$ java helper_code/HuffmanAlgorithm -< data/q32x48.bin > data/q32x48.bin.huf1
$ java helper_code/HuffmanAlgorithm + < data/q32x48.bin.huf1 | java helper_code/BinaryDump 0
1536 bits
Time taken: 2665600 ns.
```

RunLength test case (a bitmap)

Original

```
$ java helper_code/BinaryDump 0 < data/q32x48.bin
1536 bits
Time taken: 4365101 ns.
```

Compressed

```
$ java helper_code/RunLength - < data/q32x48.bin > data/q32x48.bin.rle2
$ java helper_code/BinaryDump 0 < data/q32x48.bin.rle2
1144 bits               compression ratio 1144/1536 = 74%
Time taken: 2758801 ns.
```

Decompressed

```
$ java helper_code/RunLength -< data/q32x48.bin > data/q32x48.bin.rle2
$ java helper_code/RunLength + < data/q32x48.bin.rle2 | java helper_code/BinaryDump 0
1536 bits
Time taken: 7167600 ns.
```

Binary Test case(genomeVirus)

Original

```
$ java helper_code/BinaryDump 0 < data/genomeVirus.txt
50008 bits
Time taken: 6237800 ns.
```

Compressed

```
$ java helper_code/HuffmanAlgorithm - < data/genomeVirus.txt > data/genomeVirus.txt.huf
$ java helper_code/BinaryDump 0 < data/genomeVirus.txt.huf
12576 bits        compression ratio 12576/50008 = 25%
Time taken: 9628000 ns.
```

Decompressed

```
$ java helper_code/HuffmanAlgorithm -< data/genomeVirus.txt > data/genomeVirus.txt.huf
$ java helper_code/HuffmanAlgorithm + < data/genomeVirus.txt.huf | java helper_code/BinaryDump 0
50008 bits
Time taken: 20413300 ns.
```

Hex Test case(genomeVirus)

Original

```
$ java helper_code/HexDump 0 < data/genomeVirus.txt
50008 bits
*The time taken: 60 ns
```

Compressed

```
$ java helper_code/HuffmanAlgorithm - < data/genomeVirus.txt > data/genomeVirus.txt.huf2
$ java helper_code/HexDump 0 < data/genomeVirus.txt.huf2
12576 bits        compression ratio 12576/50008 = 25%
*The time taken: 52 ns.
```

Decompressed

```
$ java helper_code/HuffmanAlgorithm -< data/genomeVirus.txt > data/genomeVirus.txt.huf2
$ java helper_code/HuffmanAlgorithm + < data/genomeVirus.txt.huf2 | java helper_code/HexDump 0
50008 bits
*The time taken: 39 ns
```

Binary Test case(mobyDick)

Original

```
$ java helper_code/BinaryDump 0 < data/mobyDick.txt
9708968 bits
Time taken: 67857300 ns.
```

Compressed

```
$ java helper_code/HuffmanAlgorithm - < data/mobyDick.txt > data/mobyDick.txt.huf
$ java helper_code/BinaryDump 0 < data/mobyDick.txt.huf
5505432 bits        compression ratio 5505432/9708968 = 57%
Time taken: 82481900 ns.
```

Decompressed

```
$ java helper_code/HuffmanAlgorithm -< data/mobyDick.txt > data/mobyDick.txt.huf
$ java helper_code/HuffmanAlgorithm + < data/mobyDick.txt.huf | java helper_code/BinaryDump 0
9708968 bits
Time taken: 138303101 ns.
```

Hex Test case(mobyDick)

Original

```
$ java helper_code/HexDump 0 < data/mobyDick.txt
9708968 bits
Time taken: 77 ns.
```

Compressed

```
$ java helper_code/HuffmanAlgorithm - < data/mobyDick.txt > data/mobyDick.txt.huf1
$ java helper_code/HexDump 0 < data/mobyDick.txt.huf1
5505432 bits        compression ratio 5505432/9708968 = 57%
Time taken: 66 ns.
```

Decompressed

```
$ java helper_code/HuffmanAlgorithm -< data/mobyDick.txt > data/mobyDick.txt.huf1
$ java helper_code/HuffmanAlgorithm + < data/mobyDick.txt.huf1 | java helper_code/HexDump 0
9708968 bits
Time taken: 117 ns.
```

Hex Test case(medTale)

Original

```
$ java helper_code/HexDump 0 < data/medTale.txt
45872 bits
Time taken: 28 ns.
```

Compressed

```
$ java helper_code/HuffmanAlgorithm - < data/medTale.txt > data/medTale.txt.huf1
$ java helper_code/HexDump 0 < data/medTale.txt.huf1
24664 bits        compression ratio 24664 /45872 = 54%
Time taken: 19 ns.
```

Decompressed

```
$ java helper_code/HuffmanAlgorithm -< data/ medTale.txt > data/medTale.txt.huf1
$ java helper_code/HuffmanAlgorithm + < data/ medTale.txt.huf1 | java helper_code/HexDump 0
45872 bits
Time taken: 57 ns.
```

Binary Test case(medTale)

Original

```
$ java helper_code/BinaryDump 0 < data/medTale.txt
45872 bits
Time taken: 6573800 ns.
```

Compressed

```
$ java helper_code/HuffmanAlgorithm - < data/ medTale.txt > data/medTale.txt.huf
$ java helper_code/BinaryDump 0 < data/medTale.txt.huf
24664 bits        compression ratio 24664 /45872 = 54%
Time taken: 10805001 ns.
```

Decompressed

```
$ java helper_code/HuffmanAlgorithm -< data/ medTale.txt > data/medTale.txt.huf
$ java helper_code/HuffmanAlgorithm + < data/ medTale.txt.huf | java helper_code/BinaryDump 0
45872 bits
Time taken: 10049801ns.
```

Bitmap test case (1536 bits 42 ns)

Compressed

```
$ java helper_code/HuffmanAlgorithm - < data/q32x48.bin > data/q32x48.bin.huf
$ java helper_code/HexDump 0 < data/q32x48.bin.huf
816 bits        compression ratio 816/1536 = 53%
Time taken: 25 ns.
```

Decompressed

```
$ java helper_code/HuffmanAlgorithm -< data/q32x48.bin > data/q32x48.bin.huf
$ java helper_code/HuffmanAlgorithm + < data/q32x48.bin.huf | java helper_code/HexDump 0
1536 bits
Time taken: 16 ns.
```

Bitmap test case (1536 bits 4365101 ns)

Compressed

```
$ java helper_code/HuffmanAlgorithm - < data/q32x48.bin > data/q32x48.bin.huf1
$ java helper_code/BinaryDump 0 < data/q32x48.bin.huf1
816 bits        compression ratio 816/1536 = 53%
Time taken: 2235399 ns.
```

Decompressed

```
$ java helper_code/HuffmanAlgorithm -< data/q32x48.bin > data/q32x48.bin.huf1
$ java helper_code/HuffmanAlgorithm + < data/q32x48.bin.huf1 | java helper_code/BinaryDump 0
1536 bits
Time taken: 2665600 ns.
```

RunLength test case (a bitmap)

Original

```
$ java helper_code/HexDump 0 < data/q32x48.bin
1536 bits
Time taken: 42 ns.
```

Compressed

```
$ java helper_code/RunLength - < data/q32x48.bin > data/q32x48.bin.rle
$ java helper_code/HexDump 0 < data/q32x48.bin.rle
1144 bits                compression ratio 1144/1536 = 74%
Time taken: 55 ns.
```

Decompressed

```
$ java helper_code/RunLength -< data/q32x48.bin > data/q32x48.bin.rle
$ java helper_code/RunLength + < data/q32x48.bin.rle | java
helper_code/HexDump 0
1536 bits
Time taken: 68 ns.
```

RunLength test case (ASCII Text)

Original

```
$ java helper_code/BinaryDump 0 < data/genomeVirus.txt
50008 bits
Time taken: 14670999 ns.
```

Compressed

```
$ java helper_code/RunLength - < data/genomeVirus.txt >
data/genomeVirus.txt.rle
$ java helper_code/BinaryDump 0 < data/ genomeVirus.txt.rle
223632 bits              compression ratio 223632/50008 = 447%...very high
Time taken: 41119801 ns.
```

Decompressed

```
$ java helper_code/RunLength -< data/ genomeVirus.txt > data/
genomeVirus.txt.rle
$ java helper_code/RunLength + < data/ genomeVirus.txt.rle | java
helper_code/BinaryDump 0
1536 bits
Time taken: 17248599 ns.
```

RunLength test case (a bitmap)

Original

```
$ java helper_code/BinaryDump 0 < data/q32x48.bin
1536 bits
Time taken: 4365101 ns.
```

Compressed

```
$ java helper_code/RunLength - < data/q32x48.bin > data/q32x48.bin.rle2
$ java helper_code/BinaryDump 0 < data/q32x48.bin.rle2
1144 bits                compression ratio 1144/1536 = 74%
Time taken: 2758801 ns.
```

Decompressed

```
$ java helper_code/RunLength -< data/q32x48.bin > data/q32x48.bin.rle2
$ java helper_code/RunLength + < data/q32x48.bin.rle2 | java helper_code/BinaryDump 0
1536 bits
Time taken: 7167600 ns.
```

RunLength test case (ASCII Text)

Original

```
$ java helper_code/BinaryDump 0 < data/medTale.txt
45872 bits
Time taken: 4937200 ns.
```

Compressed

```
$ java helper_code/RunLength - < data/medTale.txt > data/medTale.txt.rle
$ java helper_code/BinaryDump 0 < data/medTale.txt.rle
185784 bits              compression ratio 185784/45872 = 405%...very high
Time taken: 28908600 ns.
```

Decompressed

```
$ java helper_code/RunLength -< data/ genomeVirus.txt > data/ genomeVirus.txt.rle
$ java helper_code/RunLength + < data/ genomeVirus.txt.rle | java helper_code/BinaryDump 0
1536 bits
Time taken: 43 ns.
```

RunLength test case (ASCII Text)

Original

```
$ java helper_code/BinaryDump 0 < data/mobyDick.txt
9708968 bits
Time taken: 64844801 ns.
```

Compressed

```
$ java helper_code/RunLength - < data/mobyDick.txt > data/mobyDick.txt.rle
$ java helper_code/BinaryDump 0 < data/mobyDick.txt.rle
39407992 bits            compression ratio 39407992/9708968 = 406%...very high
Time taken: 210466401 ns.
```

Decompressed

```
$ java helper_code/RunLength -< data/mobyDick.txt > data/mobyDick.txt.rle
$ java helper_code/RunLength + < data/mobyDick.txt.rle | java helper_code/BinaryDump 0
9708968 bits
Time taken: 365207501 ns.
```