38.1 Introduction to Compression

As an introduction to compression, consider the processes of creating and unzipping a zip file.

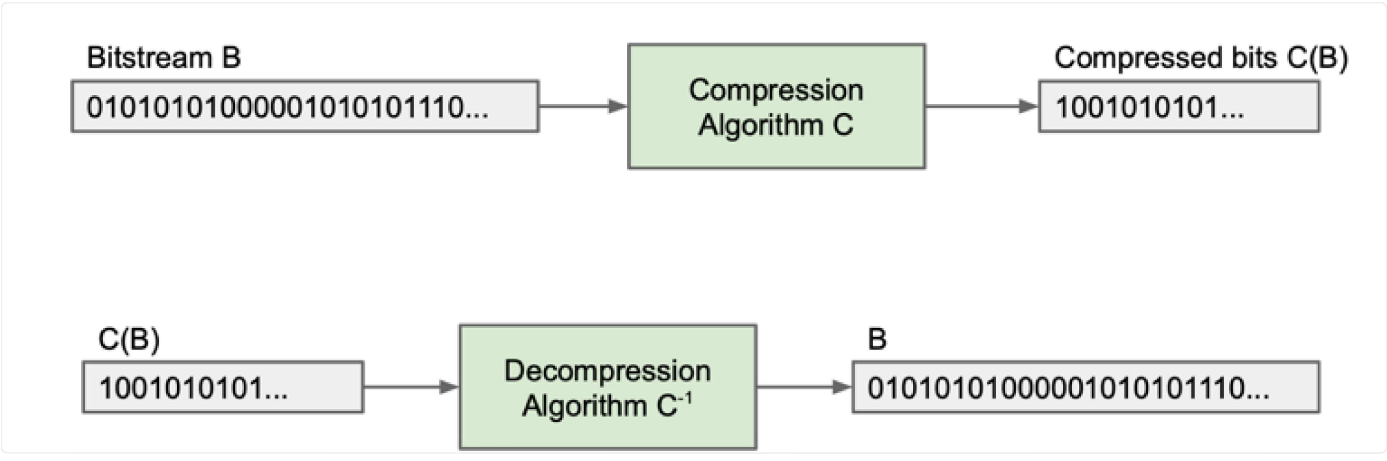
```
$ zip mobydick.zip mobydick.txt
adding: mobydick.txt (deflated 59%)

$ ls -l
-rw-rw-r-- 1 jug jug 643207 Apr 24 10:55 mobydick.txt
-rw-rw-r-- 1 jug jug 261375 Apr 24 10:55 mobydick.zip
```

Note that before and after unzipping, the file size changes!

Compression Model 1: Algorithms on Bits

In our first model of compression, we consider compression as applying a *compression algorithm* on a sequence of bits. To reverse the compression, we apply the inverse *decompression algorithm.*



Compression and decompression.

Say you had a text file called `example.txt`. If you were to zip that text file, you'd get `example.zip`, a zip file with a size much lesser than the original `example.txt` file. This is

the main idea behind compression--a technique used to reduce file size.

Then, if you were to unzip `example.zip` into a file called `unzippedexample.txt`, you would notice no difference between `example.txt` and `unzippedexample.txt.` This is an indicator of **lossless** compression, where no information is lost.

Previous
38. Compression and Complexity

Next
38.2 Prefix-free Codes

Last updated 1 year ago