CS240, Spring 2022
Assignment 5: Question 3

Q3a) What do you need to change in the RLE from the slides to make it work for a ternary string?
Since our alphabet for encoding can now contain 2, our encoding can now be expressed in base 3 rather then base 2. Since our base has increased, we will need to update Elias Gammas encoding to use log_3 rather then log_2 because we now have 3 possibilities for each possible letter (0,1,2).
As well since we now have three possibilities for our alphabet, how we pick the leading character of each encryption sequence will change. We will now check the top value in the stream rather then just $b = b - 1$

Q3b) What are your ideas for making sure you are optimizing the compression rate?
Since we are now using base 2 rather then base 3, the length of the encoding will be smaller in some cases. For example:

Base 2 of 5 : 101

Base 3 of 5 : 12

Since the number of pre appended zeros matches the length of the string, we thus will optimize the compression rate by using log_3 rather then log_2.
Q3c) Present your algorithm as a pseudocode.

Algorithm 1 Ternary RLE encoding

```
Require: S, input-stream of bits, C: output-stream
  b ← S.top()
  while S is non-empty do
    C.append(b)
    k ← 1
    while S is non-empty and S.top() == b do
      k++
      S.pop()
    end while
    K ← empty string
    while k > 2 do
      C.append(0);
      K.prepend(k mod 3)
      k ← ⌊k/3⌋
    end while
    K.prepend(k)
    C.append(K)
    b ← S.top()
  end while
```

Q3d) What is the compression rate of your encoding scheme for the following ternary string?

In order to make the string more readable we will break it into a list of pairs {a,b} where a represents the character being repeated and b represents the number of repeats. Thus the string can be represented by:

$$\{0,1\},\{1,1\},\{2,1\},\{0,2\},\{1,2\},\{2,2\},\{0,3\},\{1,3\},\{2,3\},\{0,4\},\{1,4\},\{2,4\},\{0,14\},\{1,16\},\{2,17\}$$

Plugging this into our algorithm our output is:

$$01112102122200101010201000111011201100011210012120012 2$$

Thus our compression rate will be:

$$\text{compression rate} : \frac{54}{77} \approx 70\%$$