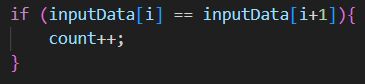
RLE – принцип сокращения веса текста, с помощью преобразования повторений в специальные блоки. Другими словами, алгоритм сжатия текста.

Из книги:

Алгоритм RLE (англ. Run-length encoding) — простой алго ритм сжатия данных, который оперирует сериями данных, т. е. последовательностями, в которых один и тот же символ встре чается несколько раз подряд. При кодировании строка одина ковых символов, составляющих серию, заменяется строкой, ко торая содержит сам повторяющийся символ и количество его повторов.

Кодирование.

Кодирование идёт только от 4 и больше символов. Если 1-3 символа - оставляем как есть.

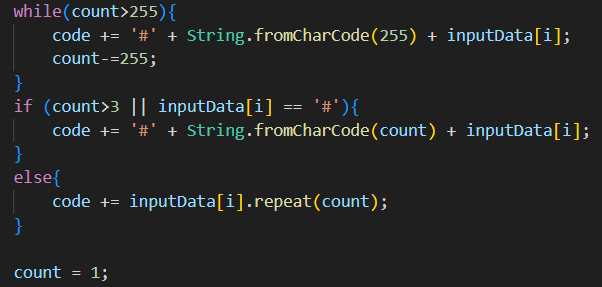


Если подряд идёт больше, чем 3 символа, то разбиваем на два процесса – до 256 и больше. Кодирование вообще происходит благодаря тому, что мы считаем кол-во повторений подряд символа, переводим это в символ с соответствующим кодом в ASCII и записываем.

Если подряд идёт от 4 до 255 символов, то мы записываем это повторение как #\*А – где # указатель на начало повторения, \* - количество символов, а А – сам символ. Если подряд идёт больше символов, чем 255, то мы просто записываем несколько раз подряд комбинации по 255 и остаток.

Так как мы используем # в качестве указателя, то любое количество “#” в тексте кодируется.

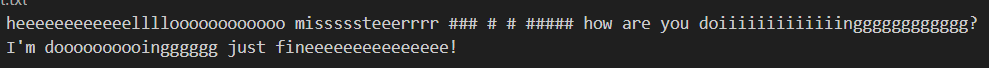
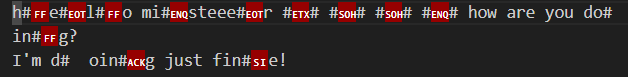
Кодируем так, чтобы не терять свободные байты. Наша сжимаемая строка занимает 3 байта (#\*А), поэтому если мы будем кодировать отрезки текста с 1-3 символами, мы либо увеличим размер, либо оставим таким же, но увеличим сложность. Мы увеличиваем размер только при кодировании строк из #, потому что иначе никак не сделать.



Пример работы:

На вход передаём действие, файл, откуда взять текст и файл, куда записать закодированный текст.

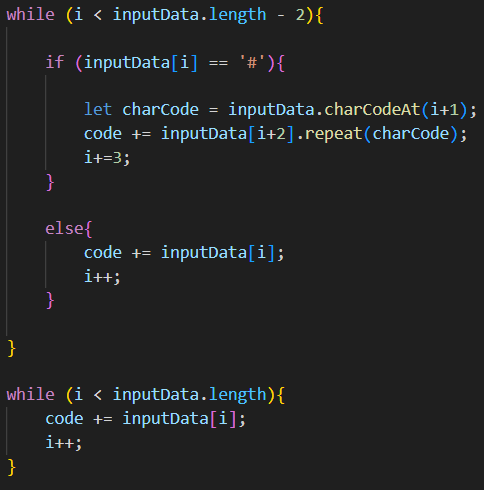


Вот так скрипт закодировал текст.

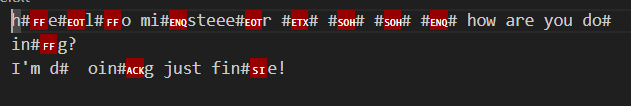
Декодирование.

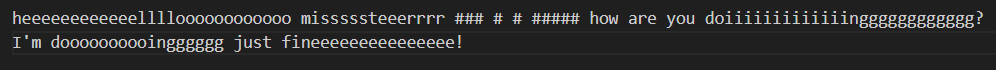
Скрипт пробегает по всему данному ему тексту, и если видит ключ #, то расшифровывает следующие два символа, как количество повторений и символ для повторения.



Пример работы:







Вывод: можем заметить, что скрипт идеально сжимает текст и разжимает.