Энтропия – алгоритм, вычисляющий энтропию строки (меру неопределённости (непредсказуемости, неожиданности) появления символа).

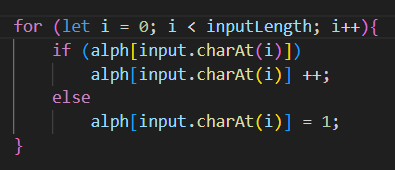
Из книги:

Энтропия — это 1) количество неопределенности относительно того, что послал источник, которое снялось после получения этого одного символа, или, что то же самое, 2) количество информации относительно того, что послал источник, которое мы получили с этим одним символом.

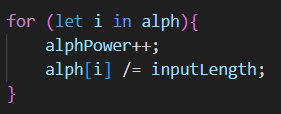
На вход в командой строке подаётся название файла, из которого происходит считывание текста, для подсчёта энтропии.



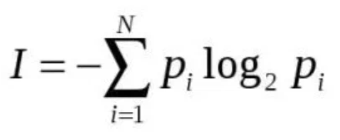
Сначала скрипт подсчитывает количество каждого отличающегося элемента.



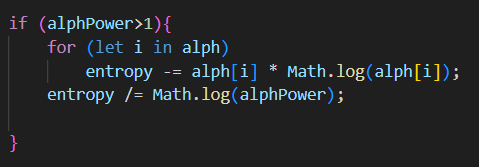
После вычисляется мощность алфавита и заодно каждое количество отдельного символа делится на общее количество символов.



Затем, если мощность алфавита больше одного, мы вычисляем энтропию. Само по себе вычисление энтропии — это формула Шеннона, но вместо 2 мы используем мощность алфавита. Pi при этом это количество определённого элемента, делённое на количество всех элементов.



Так как в js math.log() это натуральный логарифм, мы сначала делаем вычисление по формуле Шеннона, использую е как основание логарифма, а после всю получившуюся сумму делим на натуральный логарифм мощности алфавита.



При мощности алфавита меньшей или равной одному, энтропия будет равна нулю.

По итогу получим на вывод энтропию числа и таблицу с частотами каждой буквы (либо ошибку, если нет входных данных).

