Энтропия

Алгоритм вычисления энтропии текста основан на формуле **Шеннона**, которая используется в теории информации. Энтропия измеряет среднее количество информации, передаваемой на символ текста.

**Постановка задачи:**

1. На вход подается строка — последовательность букв. Требуется: построить алфавит (т. е. множество всех различных символов исходной строки) и найти частоту для всех символов алфавита.
2. Для полученной таблицы рассчитать энтропию.

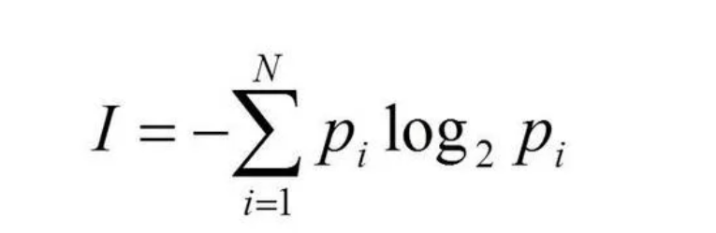
**Решение:**

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

****

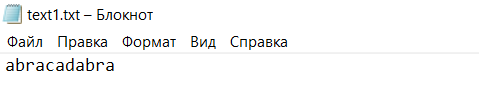
Весь код представлен в файле **entropy.js.**

Подсчет энтропии осуществляется по формуле Шеннона, но вместо 2 мы используем мощность алфавита. Pi при этом это количество определённого элемента, делённое на количество всех элементов.

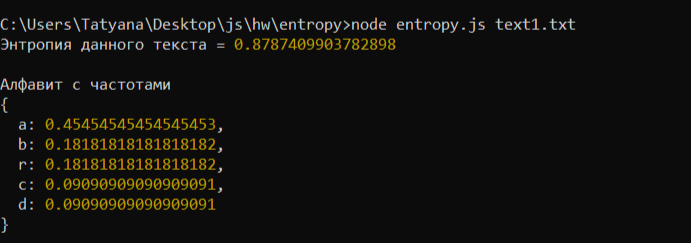


**Примеры запусков:**

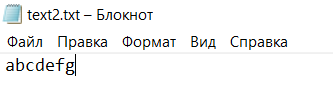
В файле text1.txt находится строка.

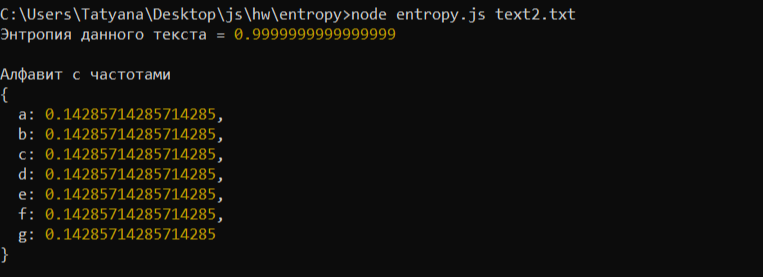


После запуска получаем:

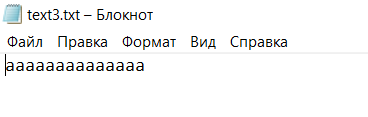


Рассмотрим случай, когда строка не содержит повторяющиеся символы (text2.txt)

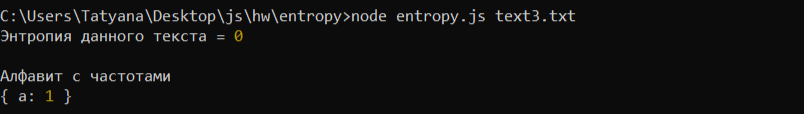


После запуска получаем:  


И рассмотрим последний пример, когда строка состоит из одной буквы (text3.txt):



Получаем:



Таким образом, данный код вычисляет энтропию текста и подсчитывает частоты символов