БЕЛОРУССКИЙ ГОСУДАРСТВЕННЫЙ ТЕХНОЛОГИЧЕСКИЙ УНИВЕРСИТЕТ

КАФЕДРА ИНФОРМАЦИОННЫХ СИСТЕМ И ТЕХНОЛОГИЙ

**Лабораторная работа №2**

**«Элементы теории информации.**

**Параметры и характеристики дискретных информационных систем**»

**Выполнила:**

студентка 3 курса 1 группы

Потапейко Полина Павловна

**Проверил:**

Берников Владислав Олегович

Минск 2021

**Цель:** приобретение практических навыков расчета и анализа параметров и информативных характеристик дискретных ИС.

**Ход работы**

**Задание а.** Рассчитать энтропию русского алфавита, определить вероятность появления каждого символа алфавита во входном тексте.

В качестве входного текста взят произвольный электронный текстовый документ на основе русского алфавита (рисунок 1).

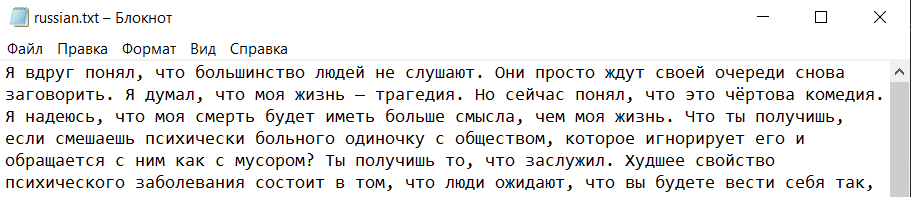


Рисунок 1 – Часть входного текста

Необходимо для дальнейшей работы перевести текст в двоичный формат. Фрагмент переведенного текста можно видеть на рисунке 2.

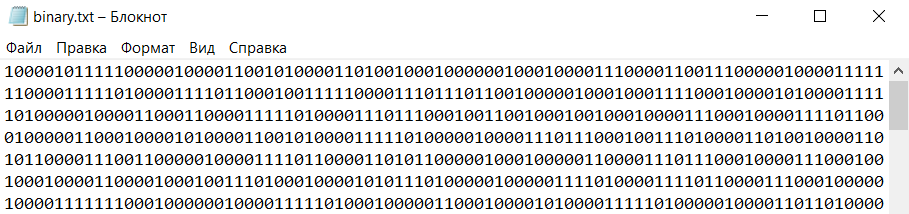


Рисунок 2 – Часть входного текста (двоичный формат)

Вероятность появления каждого символа алфавита в тексте высчитывается как частное от деления частоты появления символа, выраженной в виде целого неотрицательного числа, и общего количества всех букв в тексте.

Частоты появления символов русского алфавита можно видеть на рисунке 3.



Рисунок 3 – Частота появления символов русского алфавита

Для расчета энтропии алфавита использовалась следующая формула:

где – вероятность появления символа алфавита.

В результате вычислений энтропия русского алфавита составила 3,76954.

**Задание б.** Рассчитать энтропию бинарного алфавита для входного документа, представленного в бинарном виде.

Аналогично заданию а нужно найти вероятности появления каждого символа бинарного алфавита (рисунок 4).

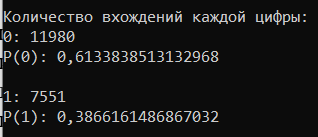


Рисунок 4 – Частота появления символов бинарного алфавита

По известной формуле энтропии нужно найти энтропию бинарного алфавита (рисунок 5).



Рисунок 5 – Энтропия бинарного алфавита

**Задание в.** Используя значения энтропии алфавитов, полученных в задании а и б, подсчитать количество информации в сообщении, состоящем из собственных фамилии, имени и отчества, на основе исходного алфавита и в кодах ASCII.

Количество информации в сообщении, состоящем из k символов, вычисляется по следующей формуле:

где – энтропия алфавита с соответствующим распределением вероятностей

Таким образом, количество информации в сообщении, состоящем из собственных фамилии, имени и отчества, посчитанном с использованием энтропии исходного алфавита представлено на рисунке 6.

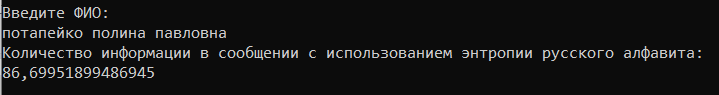


Рисунок 6 – Количество информации в сообщении (в исходном коде)

Для расчета количества информации в сообщении, представленном в коде ASCII, нужно умножить количество символов в сообщении на 8 (8 битов для одного кириллического символа), а затем умножить на значение энтропии бинарного алфавита. Результат расчета количества информации представлен на рисунке 7.



Рисунок 7 – Количество информации в сообщении (в коде ASCII)

**Задание г.** Выполнить задание пункта в при условии, что вероятность ошибочной передачи единичного бита сообщения 0.1, 0.5 и 1.0.

Так как вероятность ошибочной передачи бита сообщения больше нуля, то количество информации при его передаче будет определяться не энтропией соответствующего алфавита, а эффективной энтропией, для которой также нужно вычислить условную энтропию. Пример выполнения программы вычисления количества информации, при заданной вероятности ошибочной передачи представлен на рисунке 8.

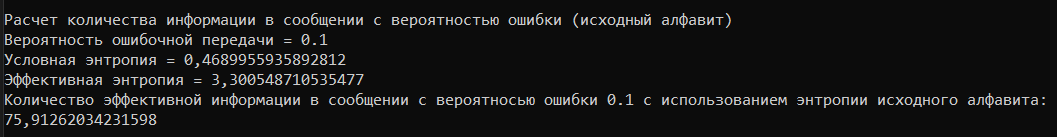


Рисунок 8 – Расчет количества информации при заданной вероятности ошибки

**Вывод:** важной информационной характеристикой дискретной информационной системы является энтропия, которая показывает, какое количество информации приходится в среднем на один символ алфавита, энтропия используется для расчета количества информации в сообщении, если сообщение передается с ошибками, то для расчета количества информации используется эффективная энтропия.