МИНИСТЕРСТВО ОБРАЗОВАНИЯ РЕСПУБЛИКИ БЕЛАРУСЬ

Учреждение образование «Белорусский государственный технологический университет»

Кафедра информационных систем и технологий

**«Отчет к лабораторной работе №2**

**Тема: Элементы теории информации. Параметры и характеристики дискретных информационных систем»**

Студент:

Септилко Анастасия Антоновна

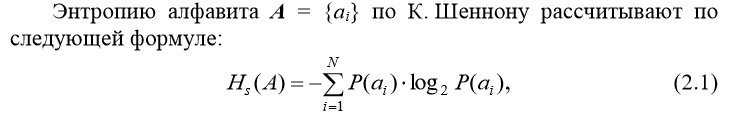
Преподаватель:

Блинова Евгения Александровна

Минск 2020

**Задания:**

**а)** рассчитать энтропию указанного преподавателем алфавитов: один – на латинице, другой – на кириллице(по формуле (2.1) – перейти от частоты появления каждого символа алфавита к соответствующей вероятности); в качестве входного может быть принят произвольный электронный текстовый документ на основе соответствующего алфавита; частоты появления символов алфавитов оформить в виде гистограмм (можно воспользоваться приложением MS Excel);



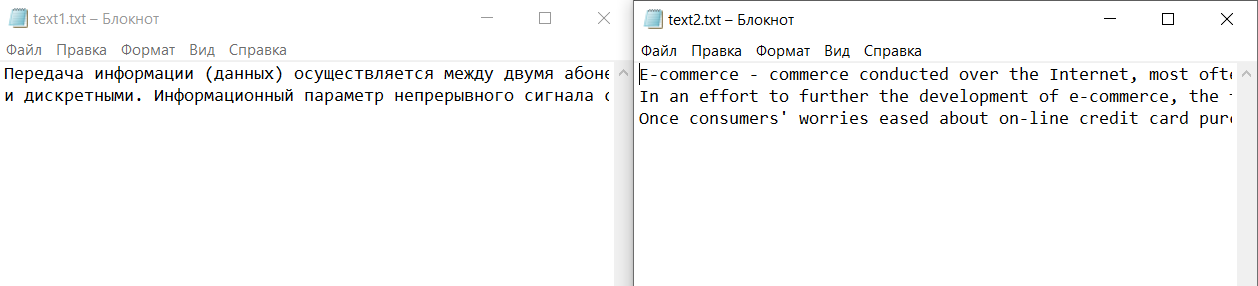
В качестве входного я создала произвольный электронный текстовый документ на основе соответствующего алфавита для русского алфавита и английского. Туда я поместила рандомный текст для подсчета энтропии (текст продемонстрирован на рисунке 1.1)

Рисунок 1.1

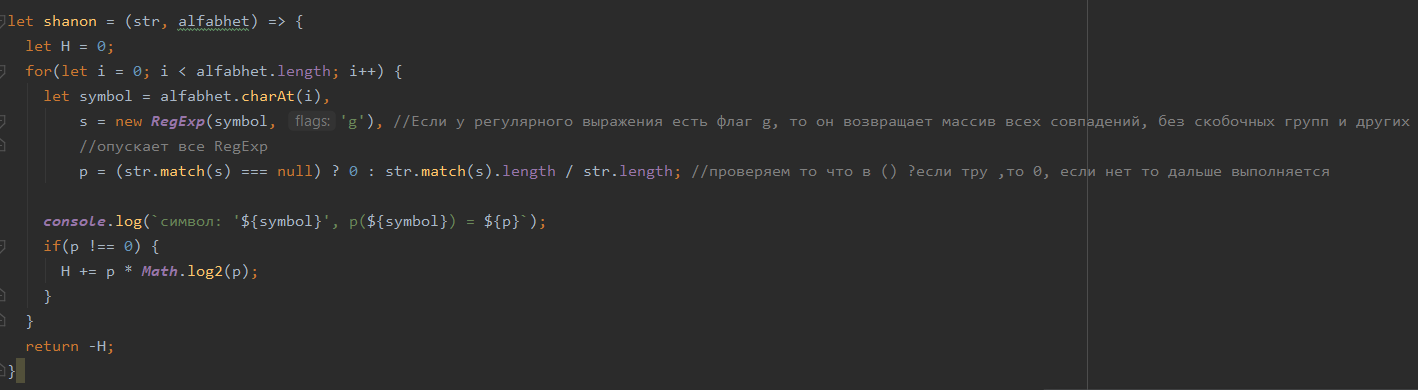
Энтропию алфавита А = {ai} по К. Шеннону рассчитывают по формуле (2.1), которая продемонстрирована выше. Здесь я привела пример (рисунок 1.2) функии для подсчета энтропии.

Рисунок 1.2

Результат выполнения программы представлен на рисунке 1.3:





Рисунок 1.3.

Частоты появления символов алфавитов оформила в виде гистограммы (воспользовалась приложением MS Excel).Пример на рисунке 1.4.



Рисунок 1.4

**б)** для входных документов, представленных в бинарных кодах, определить энтропию бинарного алфавита;

Исходный текстовый файл на рисунке 1.5:

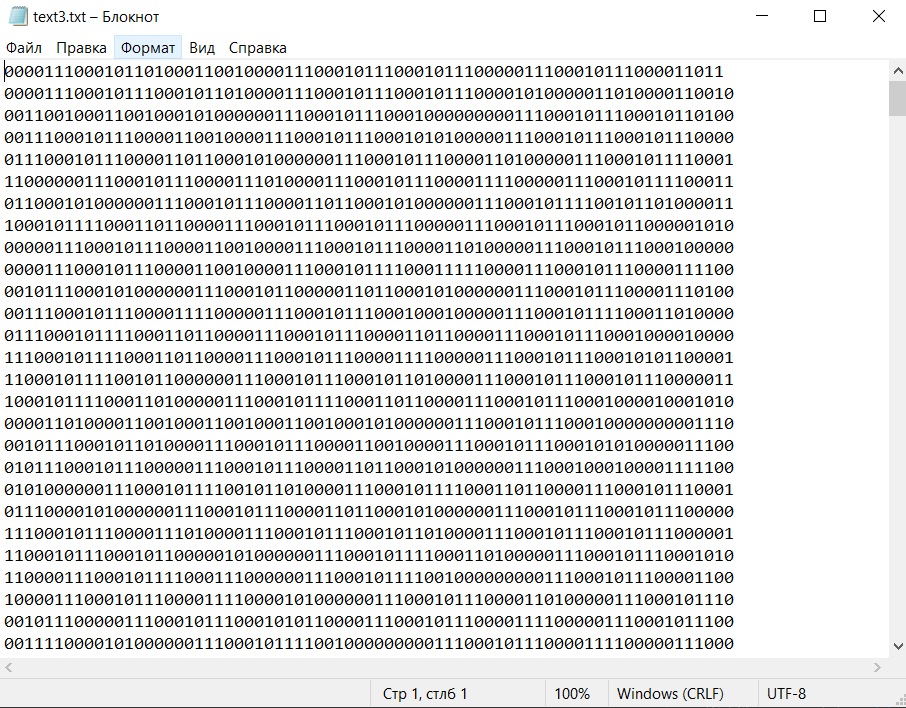


Рисунок 1.5

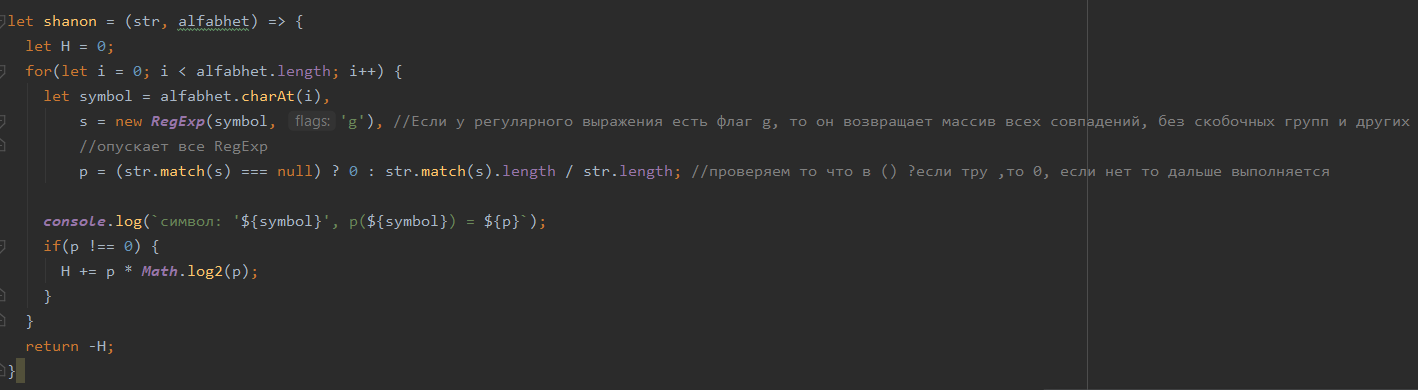
Использовали для вычисления функцию которая на рисунке 1.6:

Рисунок 1.6

Результат выполнения можно увидеть на рисунке 1.7:

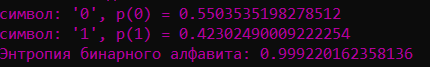


Рисунок 1.7

**в)** используя значения энтропии алфавитов, полученных в пп. а) и б), подсчитать количество информации в сообщении, состоящем из собственных фамилии, имени по отчества (на основе исходного алфавита – а) и в кодах ASCII – б); объяснить полученный результат;

Сообщение Хk, которое состоит из k символов, должно характеризоваться определенным количеством информации, I(Хk):

I(Хk) = H(A) \* k.

ЗдесьН(А)– энтропия алфавита с соответствующим распределением вероятностей р(аi)

Если принять, что р(аi =1)= р(1) и р(аi =0) = р(0), используя, вычислим энтропию бинарного алфавита:

H(A2) = -р(0)\*log2(р(0)) -р(1)\*log2(р(1)).

Функции для подсчета количества информации (рисунок 1.8):

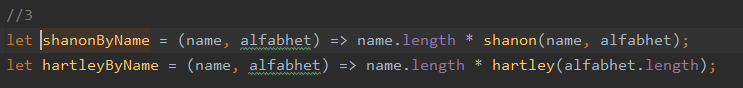


Рисунок 1.8

Строка для подсчета (рисунок 1.9):

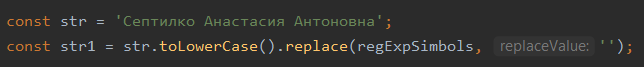


Рисунок 1.9

Вот что я получила в итоге выполнения программы (рисунок 1.10):

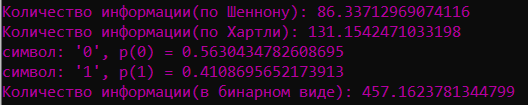


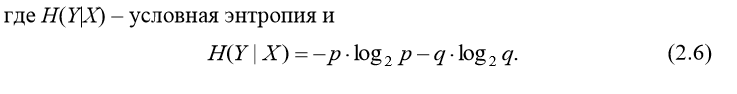
Рисунок 1.10

**г)** выполнить задание п. в) при условии, что вероятность ошибочной передачи единичного бита сообщения составляет: 0.1; 0.5; 1.0

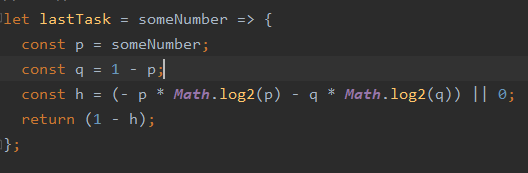
К примеру, полагая что сообщение Хk состоит только из единиц (Хk= =11…1) и имеет длину k, т.е. вероятность того, что произвольный символ равен единице, составляет единицу (р(аi =1)=1), и другая вероятность –- р(аi =0)=0 для N i , 1= . Фактически, здесь имеет место использование моноалфавита: алфавита, состоящего из одного символа. Учитывая, что сумма р(1) +р(0)==1 и, выразить одну вероятность через другую (например, р(1) = 1-р(0)), можно теоретически доказать информативность бинарного алфавита, решив дифференциальное уравнение [dH(A)/dp(1)] = 0 (вспомним из курса математики, как найти экстремум функции; можно для этого воспользоваться).

Если вероятность ошибки в ДСК отлична от 0 (р > 0), переданное сообщение может содержать ошибки: Хk ≠ Yk. Количество информации в таком сообщении при его передаче по ДСК будет определять не энтропией двоичного алфавита, а эффективной энтропией Hе(A) алфавита или пропускной способностью канала:





Продемонстрирована на рисунке 1.11 данная функция:



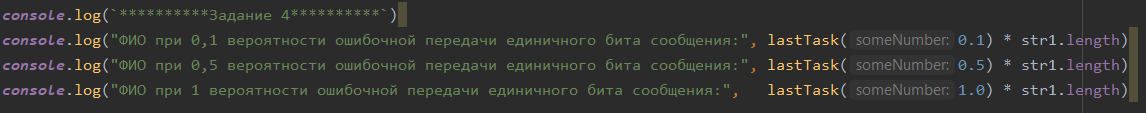


Рисунок 1.11

Итог запуска программы (рисунок 1.12):

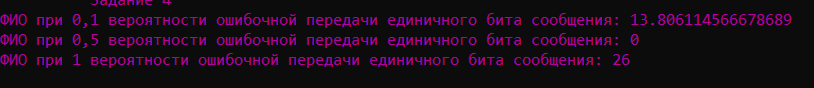


Рисунок 1.12

**Вывод**

Приобрела практические навыки расчета и анализа параметров и информативных характеристик дискретных ИС. Закрепила теоретические знания по основам теории информации. Разработала приложение для расчета и анализа параметров и информативных характеристик дискретных ИС. Результаты выполнения лабораторной работы оформила в виде описания разработанного приложения, методики выполнения экспериментов с использованием приложения и результатов эксперимента.

**Контрольные вопросы:**

**1. Что такое алфавит источника сообщения?**

Алфавит источника сообщения – это конечная или бесконечная совокупность символов и знаков, формирующих сообщение.

**2. Что такое мощность алфавита источника сообщения?**

Мощность алфавита – это количество символов, составляющих алфавит.

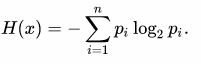
**3. Что такое энтропия алфавита?**

Энтропия алфавита – это информационная характеристика алфавита. Энтропия характеризует количество информации, приходящееся в среднем на один символ алфавита.

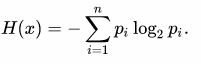
**4. От чего зависит энтропия алфавита?**

Зависит только от распределения вероятностей, а не от алфавита

**5. Напишите формулу вычисления энтропии алфавита.**



**6. Как рассчитывается энтропия по Шеннону?**



где pi – вероятность i-ого события.

**7. Как рассчитывается энтропия по Хартли?**

*,*

где N – мощность алфавита.

**8. В чем различие между энтропией по Шеннону и энтропией по Хартли?**

Энтропия по Хартли рассчитывается с учетом того, что алфавит состоит из символов с равными вероятностями появления. Для расчета энтропии по Шеннону необходимо вычислить отдельно вероятность появления каждого символа алфавита.

**9. Что такое количество информации?**

Количество информации – мера уменьшения неопределённости знаний при получении информационных сообщений.

**10. Что такое избыточность сообщений?**

Избыточность – термин, означающий превышение количества информации, используемой для передачи или хранения сообщения, над его информационной энтропией.