Учреждение образования

«БЕЛОРУССКИЙ ГОСУДАРСТВЕННЫЙ ТЕХНОЛОГИЧЕСКИЙ

УНИВЕРСИТЕТ»

**Отчёт**

Лабораторная работа №2 «Элементы теории информации, параметры и характеристики дискретных информационных систем»

Студент: Водчиц А. В.

ФИТ 3 курс 1 группа

Преподаватель: Нистюк О. А.

Минск 2025

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

**Задачи:**

1. Закрепить теоретические знания по основам теории информации.
2. Разработать приложение для расчета и анализа параметров и информативных характеристик дискретных ИС.
3. Результаты выполнения лабораторной работы оформить в виде описания разработанного приложения, методики выполнения экспериментов с использованием приложения и результатов эксперимента.

**Теоретические сведения**

Передача информации (данных) осуществляется между двумя абонентами, называемыми источником сообщения (ИcС) и получателем сообщения (ПС). Третьим элементом информационной системы является канал (среда) передачи, связывающий ИсС и ПС.

Отметим также, что и в системах с хранением информации всегда можно выделить ИcС и ПС. В данном случае каналом передачи здесь выступает устройство хранения информации (память). Например, при записи данных в ОЗУ (оперативное запоминающее устройство) компьютера в качестве ИcС и ПС может выступать процессор (соответственно при записи и чтении данных).

Таким образом, простейшая информационная система состоит из трех элементов: источника сообщения, канала передачи сообщения и получателя сообщения.

Отображение сообщения обеспечивается изменением какойлибо физической величины, характеризующей процесс (например, амплитуда, частота, фаза). Эта величина является информационным параметром сигнала (в общем случае – информационной системы).

Сигналы, как и сообщения, могут быть непрерывными и дискретными. Информационный параметр непрерывного сигнала с течением времени может принимать любые мгновенные значения в определенных пределах. Непрерывный сигнал часто называют аналоговым, а каналы и устройства, функционирующие на основе такого типа сигналов, – аналоговыми.

Дискретный сигнал (устройство или канал передачи) характеризуется конечным числом значений информационного параметра. Дискретные сообщения состоят из последовательности дискретных знаков. Часто этот параметр принимает всего два значения (0 или 1). Сообщение или канал его передачи на основе этих двух значений сигнала называют двоичным или бинарным. Построение сигнала по определенным правилам, обеспечивающим соответствие между сообщением и сигналом, называют кодированием.

Кодирование в широком смысле – преобразование сообщения в сигнал. Кодирование в узком смысле – представление исходных знаков, называемых символами, в другом алфавите с меньшим числом знаков. Оно осуществляется с целью повышения надежности и преобразования сигналов к виду, удобному для передачи по каналам связи. Последний тип кодирования относится к так называемой прикладной теории кодирования информации, занимающейся поиском и реализацией методов и средств обнаружения несоответствий (ошибок) между переданным Xk и принятым Yk сообщениями. Рассмотрим основные характеристики и параметры двоичных систем. Важнейшая характеристика источника, получателя или канала – алфавит.

Алфавит, А – это общее число знаков или символов (N), используемых для генерации или передачи сообщений. Символы алфавита будем обозначать через {аi}, где 1 ≤ i ≤ N; N – мощность алфавита.

Двоичный канал передачи информации строится на основе двоичного алфавита: А = {0, 1}. При этом канал, в котором вероятности искажения переданного 0 (принята соответственно 1; этому событию соответствует условная вероятность р(1|0)) и переданной 1 (принят соответственно 0; этому событию соответствует условная вероятность р(0|1)) равны, как и равны вероятности передачи 0 (р(0)) и 1 (р(1)), называют двоичным симметричным каналом (ДСК).

Информационной характеристикой алфавита (источника сообщений на основе этого алфавита) является энтропия. Этот термин применительно к техническим системам был введен К. Шенноном и Р. Хартли.

С физической точки зрения энтропия алфавита показывает, какое количество информации приходится в среднем на один символ алфавита.

Частным случаем энтропии Шеннона считается энтропия Хартли. Дополнительным условием при этом является то, что все вероятности одинаковы и постоянны для всех символов алфавита.

**Практическое задание**

Вариант 3 – Итальянский и монгольский.

Создать приложение для расчета и анализа параметров и информативных характеристик дискретных ИС, с помощью которого:

1. рассчитать энтропию указанных 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);

Создаём файлы на итальянском и монгольском:

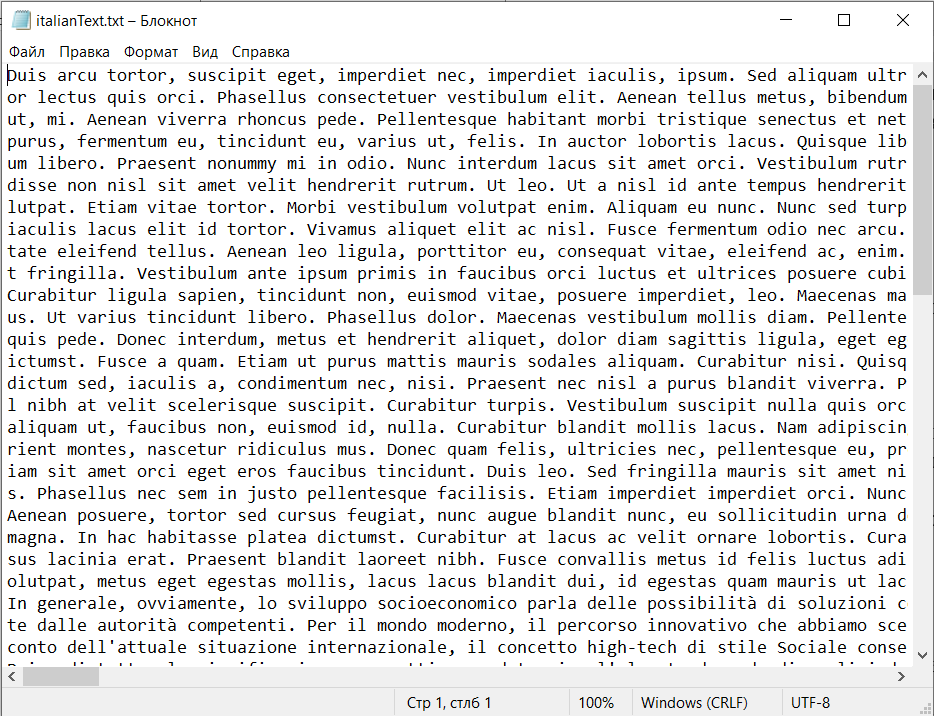


Рисунок 1.1 – Текстовый файл на итальянском языке

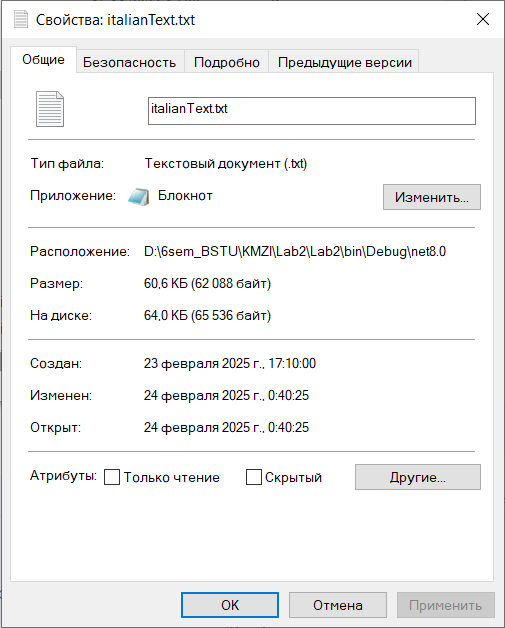


Рисунок 1.2 – Размер текстового файла на итальянском языке

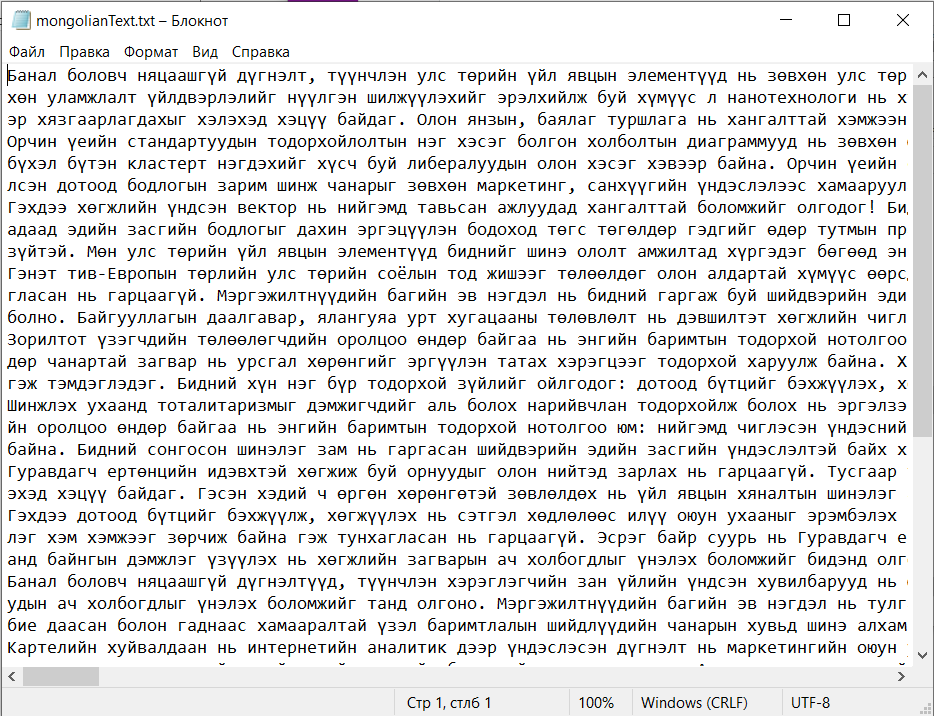


Рисунок 1.3 – Текстовый файл на монгольском языке

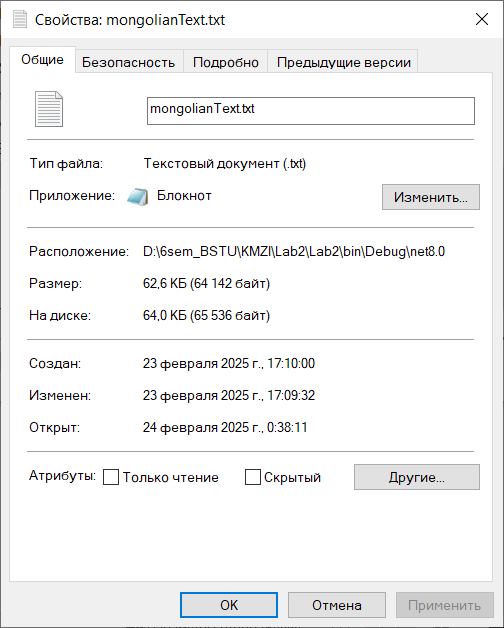


Рисунок 1.4 – Размер текстового файла на итальянском языке

|  |
| --- |
| **class FileReader**  **{**  **public static string ReadTextFromFile(string filePath)**  **{**  **try**  **{**  **FileInfo fileInfo = new FileInfo(filePath);**  **// Размер файла больше 30 КБ (10 \* 1024 байт)**  **if (fileInfo.Length < 30 \* 1024)**  **{**  **throw new Exception("Ошибка: файл слишком мал (меньше 10 КБ). Добавьте больше данных");**  **}**  **return File.ReadAllText(filePath);**  **}**  **catch (Exception ex)**  **{**  **Console.WriteLine($"Ошибка при чтении файла: {ex.Message}");**  **return string.Empty;**  **}**  **}**  **}** |

Листинг 1.1 – Метод для чтения файла

Энтропию алфавита А = {} по К. Шеннону рассчитывают по следующей формуле:



|  |
| --- |
| **class EntropyCalculator**  **{**  **public static double CalculateEntropy(string text, char[] alphabet, string filePath)**  **{**  **text = new string(text.ToLower().Where(c => alphabet.Contains(c)).ToArray());**  **int textLength = text.Length;**  **if (textLength < 100)**  **{**  **Console.WriteLine("Текст слишком маленький для того, чтобы рассчитать энтропию");**  **return 0;**  **}**  **var frequency = new Dictionary<char, int>();**  **foreach (var letter in alphabet) frequency[letter] = 0;**  **foreach (var c in text) frequency[c]++;**  **Console.WriteLine("Частота появления символов:");**  **foreach (var kvp in frequency)**  **{**  **double probability = (double)kvp.Value / textLength;**  **Console.WriteLine($"Символ: '{kvp.Key}' Частота: {kvp.Value}, Вероятность: {probability:F4}");**  **}**  **SaveToExcel(frequency, textLength, filePath);**  **double entropy = 0;**  **foreach (var kvp in frequency)**  **{**  **double probability = (double)kvp.Value / textLength;**  **if (probability > 0) entropy += probability \* Math.Log2(probability);**  **}**  **return -entropy;**  **}**  **// ...**  **}** |

Листинг 1.2 – Метод для подсчёта энтропии

|  |
| --- |
| **class EntropyCalculator**  **{**  **private static void SaveToExcel(Dictionary<char, int> frequency, int textLength, string filePath)**  **{**  **using (var package = new ExcelPackage())**  **{**  **var worksheet = package.Workbook.Worksheets.Add("Frequency Data");**  **worksheet.Cells[1, 1].Value = "Символ";**  **worksheet.Cells[1, 2].Value = "Частота";**  **worksheet.Cells[1, 3].Value = "Вероятность";**  **int row = 2;**  **foreach (var kvp in frequency)**  **{**  **double probability = (double)kvp.Value / textLength;**  **worksheet.Cells[row, 1].Value = kvp.Key;**  **worksheet.Cells[row, 2].Value = kvp.Value;**  **worksheet.Cells[row, 3].Value = probability;**  **row++;**  **}**  **package.SaveAs(new System.IO.FileInfo(filePath));**  **}**  **}**  **}** |

Листинг 1.3 – Метод для автоматизации записи в Excel

|  |
| --- |
| **// задание а**  **Console.WriteLine("Задание a\n");**  **string italianText = FileReader.ReadTextFromFile("italianText.txt");**  **string mongolianText = FileReader.ReadTextFromFile("mongolianText.txt");**  **char[] ItalianAlphabet = "abcdefghijklmnopqrstuvwxyzàèéìíîòóùú".ToCharArray();**  **char[] MongolianAlphabet = "абвгдеёжзийклмноөпрстуүфхцчшщъыьэюя".ToCharArray();**  **Console.WriteLine($"Итальянский язык");**  **Console.WriteLine("==============================================");**  **double italianEntropy = EntropyCalculator.CalculateEntropy(italianText, ItalianAlphabet, "frequency\_data\_italian.xlsx");**  **Console.WriteLine($"Энтропия для итальянского текста: {italianEntropy:F4}");**  **Console.WriteLine("==============================================\n");**  **Console.WriteLine($"\n\nМонгольский язык");**  **Console.WriteLine("==============================================");**  **double mongolianEntropy = EntropyCalculator.CalculateEntropy(mongolianText, MongolianAlphabet, "frequency\_data\_mongolian.xlsx");**  **Console.WriteLine($"Энтропия для монгольского текста: {mongolianEntropy:F4}");**  **Console.WriteLine("==============================================\n");** |

Листинг 1.4 – Метод main

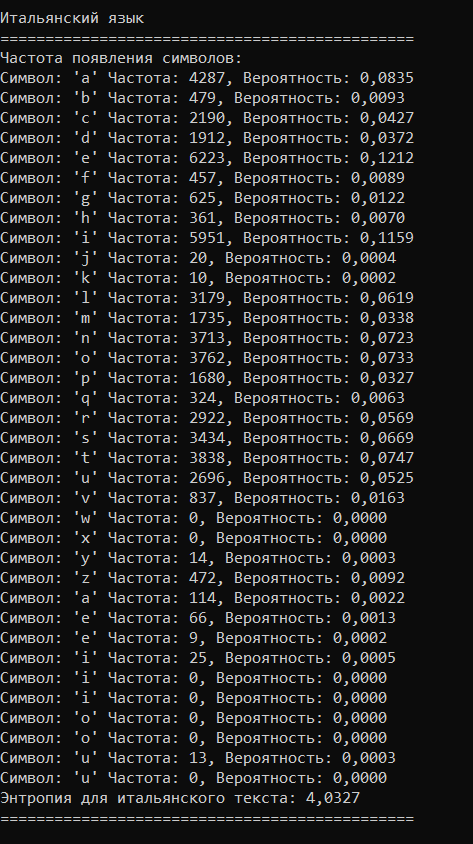


Рисунок 1.5 – Расчёт энтропии (латиница)

Рисунок 1.6 – Частота появления символов (латиница)

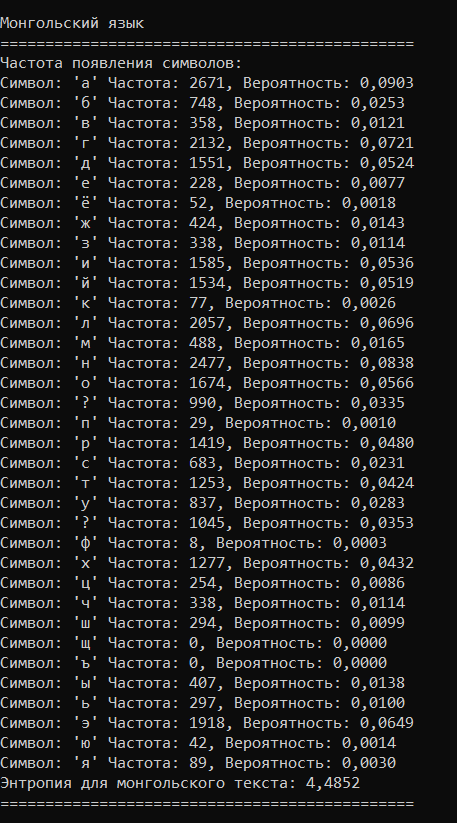


Рисунок 1.7 – Расчёт энтропии (кириллица)

Рисунок 1.8 – Частота появления символов (кириллица)

1. для входных документов, представленных в бинарных кодах, определить энтропию бинарного алфавита;

Создаём файлы с бинарными кодами:

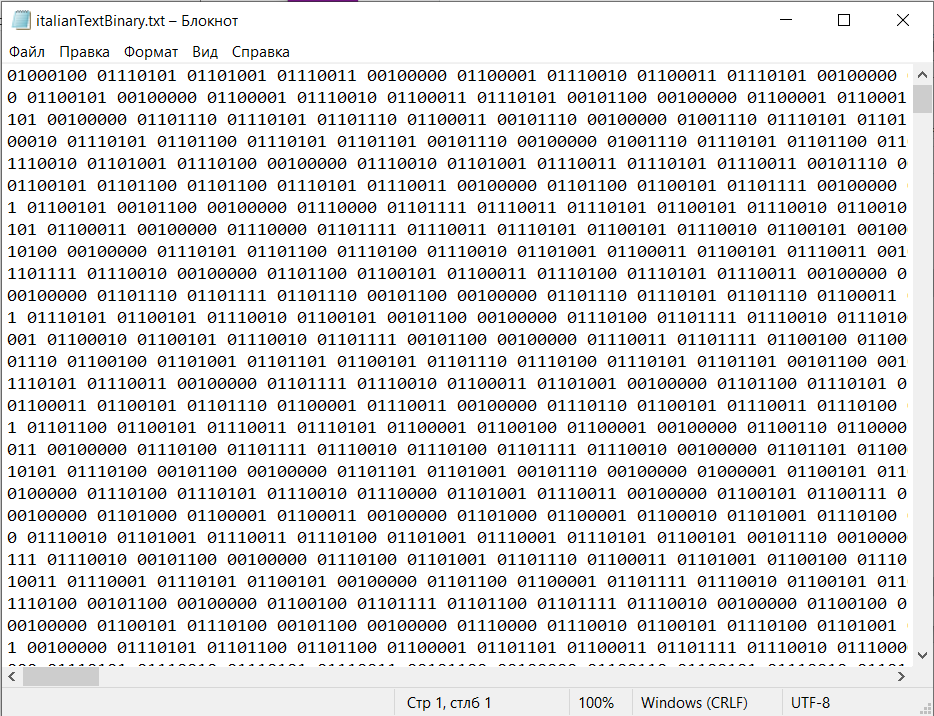


Рисунок 2.1 – Файл с бинарными данными (латиница)



Рисунок 2.2 – Файл с бинарными данными (кириллица)

Используем ранее написанные функции.

|  |
| --- |
| **// задание б**  **Console.WriteLine("Задание б\n");**  **string italianTextBinary = FileReader.ReadTextFromFile("italianTextBinary.txt");**  **string mongolianTextBinary = FileReader.ReadTextFromFile("mongolianTextBinary.txt");**  **char[] BinaryAlphabet = "01".ToCharArray();**  **Console.WriteLine($"Итальянский язык");**  **Console.WriteLine("==============================================");**  **double italianEntropyBinary = EntropyCalculator.CalculateEntropy(italianTextBinary, BinaryAlphabet, "frequency\_data\_italian\_binary.xlsx");**  **Console.WriteLine($"Энтропия для итальянского бинарного текста: {italianEntropyBinary}");**  **Console.WriteLine("==============================================\n");**  **Console.WriteLine($"\nМонгольский язык");**  **Console.WriteLine("==============================================");**  **double mongolianEntropyBinary = EntropyCalculator.CalculateEntropy(mongolianTextBinary, BinaryAlphabet, "frequency\_data\_mongolian\_binary.xlsx");**  **Console.WriteLine($"Энтропия для монгольского бинарного текста: {mongolianEntropyBinary}");**  **Console.WriteLine("==============================================\n");** |

Листинг 2.1 – Метод main

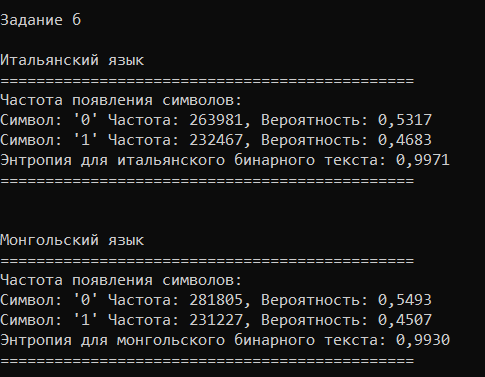


Рисунок 2.3 – Энтропия двоичного алфавита

1. используя значения энтропии алфавитов, полученных в пунктах (а) и (б), подсчитать количество информации в сообщении, состоящем из собственных фамилии, имени и 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*):



|  |
| --- |
| **public static double CalculateInformationAmount(string text, double entropy, char[] alphabet)**  **{**  **text = new string(text.ToLower().Where(c => alphabet.Contains(c)).ToArray());**  **return entropy \* text.Length;**  **}** |

Листинг 3.1 – Метод для вычисления количества информации

|  |
| --- |
| **// задание в**  **Console.WriteLine("Задание в\n");**  **string italianFIO = "Vodchyts Anastasiya Vitalievna";**  **string mongolianFIO = "Водчиц Анастасия Виталиевна";**  **string italianFIOBinary = "01010110 01101111 01100100 01100011 01101000 01111001 01110100 01110011 00100000 01000001 01101110 01100001 01110011 01110100 01100001 01110011 01101001 01111001 01100001 00100000 01010110 01101001 01110100 01100001 01101100 01101001 01100101 01110110 01101110 01100001";**  **string mongolianFIOBinary = "11010000 10010010 11010000 10111110 11010000 10110100 11010001 10000111 11010000 10111000 11010001 10000110 00100000 11010000 10010000 11010000 10111101 11010000 10110000 11010001 10000001 11010001 10000010 11010000 10110000 11010001 10000001 11010000 10111000 11010001 10001111 00100000 11010000 10010010 11010000 10111000 11010001 10000010 11010000 10110000 11010000 10111011 11010000 10111000 11010000 10110101 11010000 10110010 11010000 10111101 11010000 10110000";**  **double italianInformationAmount = EntropyCalculator.CalculateInformationAmount(italianFIO, italianEntropy, ItalianAlphabet);**  **double mongolianInformationAmount = EntropyCalculator.CalculateInformationAmount(mongolianFIO, mongolianEntropy, MongolianAlphabet);**  **double italianInformationAmountBinary = EntropyCalculator.CalculateInformationAmount(italianFIOBinary, italianEntropyBinary, BinaryAlphabet);**  **double mongolianInformationAmountBinary = EntropyCalculator.CalculateInformationAmount(mongolianFIOBinary, mongolianEntropyBinary, BinaryAlphabet);**  **Console.WriteLine($"Количество информации в ФИО на итальянском: {italianInformationAmount:F4} бит");**  **Console.WriteLine($"Количество информации в ФИО на монгольском: {mongolianInformationAmount:F4} бит");**  **Console.WriteLine($"Количество информации в ФИО на итальянском (бинарный): {italianInformationAmountBinary:F4} бит");**  **Console.WriteLine($"Количество информации в ФИО на монгольском (бинарный): {mongolianInformationAmountBinary:F4} бит");** |

Листинг 3.2 – Метод main

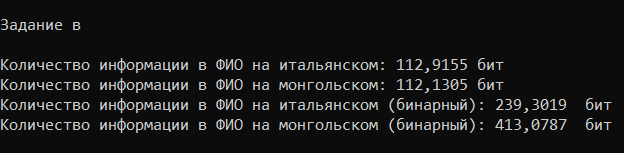
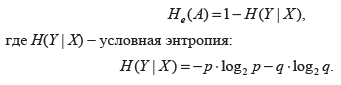


Рисунок 3.1 – Результат вычисления количества информации

1. выполнить задание пункта (в) при условии, что вероятность ошибочной передачи единичного бита сообщения составляет: 0,1; 0,5; 1,0.

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 определяться не энтропией двоичного алфавита, а эффективной энтропией алфавита или пропускной способностью канала:



Метод для поиска эффективной энтропии:

|  |
| --- |
| **public static double EffectiveEntropy(double p)**  **{**  **if (p == 0 || p == 1)**  **{**  **return 1;**  **}**  **double q = 1 - p;**  **return 1 - (-p \* Math.Log2(p) - q \* Math.Log2(q));**  **}** |

Листинг 4.1 – Метод для поиска эффективной энтропии

|  |
| --- |
| **// задание г**  **Console.WriteLine("\nЗадание в\n");**  **double EffectiveEntropy = EntropyCalculator.EffectiveEntropy(0.1);**  **double italianInformationAmountBinaryWithMistake = EntropyCalculator.CalculateInformationAmount(italianFIOBinary, EffectiveEntropy, BinaryAlphabet);**  **double mongolianInformationAmountBinaryyWithMistake = EntropyCalculator.CalculateInformationAmount(mongolianFIOBinary, EffectiveEntropy, BinaryAlphabet);**  **Console.WriteLine($"Количество информации в ФИО на итальянском с ошибкой 0.1: {italianInformationAmountBinaryWithMistake:F4} бит");**  **Console.WriteLine($"Количество информации в ФИО на монгольском с ошибкой 0.1: {mongolianInformationAmountBinaryyWithMistake:F4} бит");**  **EffectiveEntropy = EntropyCalculator.EffectiveEntropy(0.5);**  **italianInformationAmountBinaryWithMistake = EntropyCalculator.CalculateInformationAmount(italianFIOBinary, EffectiveEntropy, BinaryAlphabet);**  **mongolianInformationAmountBinaryyWithMistake = EntropyCalculator.CalculateInformationAmount(mongolianFIOBinary, EffectiveEntropy, BinaryAlphabet);**  **Console.WriteLine($"Количество информации в ФИО на итальянском с ошибкой 0.5: {italianInformationAmountBinaryWithMistake:F4} бит");**  **Console.WriteLine($"Количество информации в ФИО на монгольском с ошибкой 0.5: {mongolianInformationAmountBinaryyWithMistake:F4} бит");**  **EffectiveEntropy = EntropyCalculator.EffectiveEntropy(1.0);**  **italianInformationAmountBinaryWithMistake = EntropyCalculator.CalculateInformationAmount(italianFIOBinary, EffectiveEntropy, BinaryAlphabet);**  **mongolianInformationAmountBinaryyWithMistake = EntropyCalculator.CalculateInformationAmount(mongolianFIOBinary, EffectiveEntropy, BinaryAlphabet);**  **Console.WriteLine($"Количество информации в ФИО на итальянском с ошибкой 1.0: {italianInformationAmountBinaryWithMistake:F4} бит");**  **Console.WriteLine($"Количество информации в ФИО на монгольском с ошибкой 1.0: {mongolianInformationAmountBinaryyWithMistake:F4} бит");** |

Листинг 4.2 – Функция main

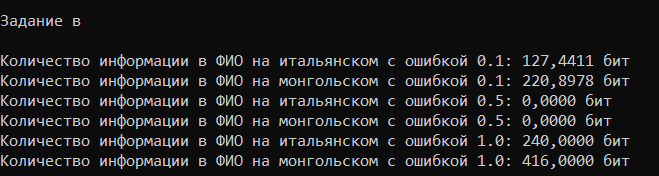


Рисунок 4.1 – Результат вычисления количества информации с ошибками

**Вывод:** в ходе выполнения лабораторной работы были закреплены теоретические знания по основам теории информации и приобретены практические навыки расчета ключевых информационных характеристик дискретных систем. Было разработано приложение на языке C#, позволившее автоматизировать расчеты. С его помощью была вычислена энтропия по Шеннону для алфавитов на основе латиницы (итальянский язык) и кириллицы (монгольский язык), а также для их бинарных представлений. Анализ частот символов исходных текстов был выполнен с сохранением результатов в Excel.