МИНИСТЕРСТВО ОБРАЗОВАНИЯ РЕСПУБЛИКИ БЕЛАРУСЬ

Учреждение образования «БЕЛОРУССКИЙ ГОСУДАРСТВЕННЫЙ

ТЕХНОЛОГИЧЕСКИЙ УНИВЕРСИТЕТ»

Отчет по лабораторной работе №2

Элементы теории информации. Параметры и характеристики дискретных информационных систем

Выполнил студент: Плюто Э. В.

ФИТ 3 курса, 5 группа

Проверил: Савельева М. Г.

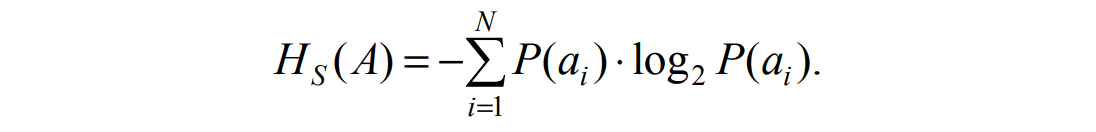
Минск 2024

**Практическое задание:**

**А).** Рассчитать энтропию указанных преподавателем алфавитов: один – на латинице, другой – на кириллице; в качестве входного может быть принят произвольный электронный текстовый документ на основе соответствующего алфавита; частоты появления символов алфавитов оформить в виде гистограмм.

Энтропия – информационная характеристика алфавита (источника сообщений на основе этого алфавита), которая показывает какое количество информации приходится в среднем на один символ алфавита.

Энтропия алфавита рассчитывается по К. Шеннону по формуле:

(1.1)

Реализация функции для расчета энтропии алфавитов на латинице (чешском) и кириллице (белорусском) представлена на листинге 1.

|  |
| --- |
| public static double GetEntropy(string text, string lang)  {  double entropy = 0;  double[] probability = new double[lang.Length];  for(int i = 0; i < lang.Length; i++)  {  probability[i] = (double)text.Count(x => x == lang[i]) / (double)text.Length;  if (probability[i] != 0)  {  Console.WriteLine(lang[i] + " - " + probability[i]);  entropy -= probability[i] \* Math.Log2(probability[i]);  }  }  return entropy;  } |

Листинг 1 – Реализация функции подсчета энтропии

Результаты выполнения вышеописанной функции представлены на рисунке 1.



Рисунок 1 – Результат выполнения функции

Гистограмма частоты появления символов в чешском алфавите представлена на рисунке 2.

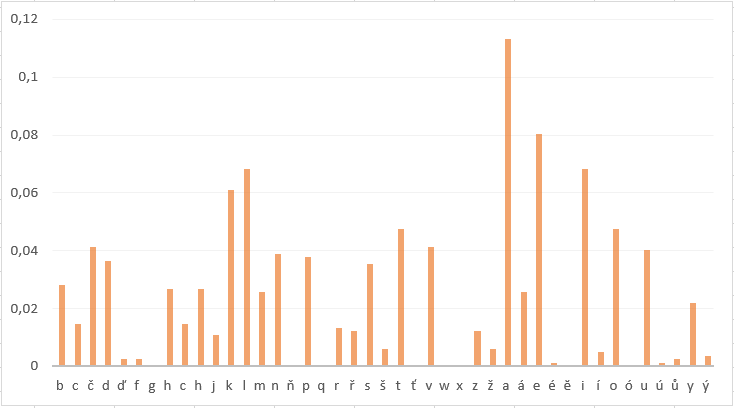


Рисунок 2 – Гистограмма чешского алфавита

Аналогичная гистограмма для белорусского алфавита представлена на рисунке 3.

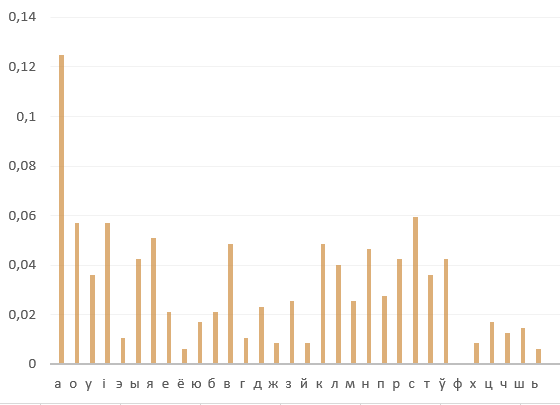


Рисунок 3 – Гистограмма белорусского алфавита

Из данных гистограмм можно сделать вывод, что наиболее часто встречаемый символ в чешском языке – «a», в белорусском - «a».

**Б).** Для входных документов, представленных в бинарных кодах, определить энтропию бинарного алфавита.

Для вычисления энтропии бинарного алфавита использовалась функция GetEntropy() из предыдущего задания.

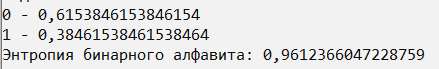


Рисунок 4 – Результат выполнения функции

**В).** 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)).

Сообщение должно характеризоваться определенным количеством информации, которое можно вычислить по формуле:



(1.2)

Результаты вычисления на основе двух алфавитов представлены на рисунке 5.



Рисунок 5 – Количество информации

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

Для подсчета количества информации с вероятностью ошибочной передачи единичного бита была реализована функция GetEffectiveEntropy(), считающая эффективную энтропию, листинг представлен ниже.

|  |
| --- |
| public static double GetEffectiveEntropy(double p, string text)  {  if (!IsBinary(text)) return 0;  if (p == 1) return 1;  double q = 1 - p;  return 1 - (-1) \* (p \* Math.Log2(p) + q \* Math.Log2(q));  } |

Листинг 2 – Реализация функции подсчета эффективной энтропии

Результат выполнения данной функции представлен на рисунке 6.

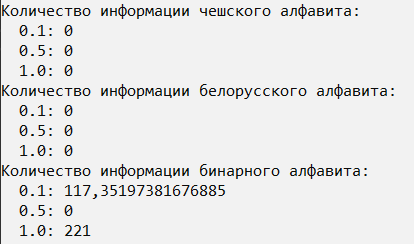


Рисунок 6 – Количество информации алфавитов

**Вывод:**

Были приобретены практические навыки расчета и анализа параметров и информативных характеристик дискретных ИС. Реализованы программы для подсчета энтропии, количества информации, количества информации с вероятностью ошибочной передачи различных алфавитов.