**Лабораторная работа №1**

***Вычисление энтропии Шеннона***

Порядок выполнения работы

1. Изучить теоретический материал гл. 2.
2. Реализовать процедуру вычисления энтропии для текстового файла на английском языке. В процедуре необходимо подсчитывать частоты появления символов (прописные и заглавные буквы не отличаются, знаки препинания  рассматриваются как один символ, пробел является самостоятельным символом), которые можно использовать как оценки вероятностей появления символов. Затем вычислить величину энтропии Шеннона. Точность вычисления -- 4 знака после запятой. Обязательно предусмотреть возможность ввода имени файла, для которого будет вычисляться  энтропия.
3. Проверить запрограммированную процедуру на нескольких файлах (размер не менее 1 Кб) и заполнить таблицу следующего вида вычисленными значениями энтропии

|  |  |  |  |
| --- | --- | --- | --- |
| Используются частоты одиночных символов | | | |
| Энтропия текста  на английском языке | Макс. значение энтропии текста  на английском языке | Энтропия разработанной программы для л.р.1 | Макс. значение энтропии разработанной программы для л.р.1 |
|  |  |  |  |

4. Вычислить значение энтропии для тех же файлов, но с использованием частот вхождений пар символов.

|  |  |  |  |
| --- | --- | --- | --- |
| Используются частоты пар символов | | | |
| Энтропия текста  на английском языке | Макс. значение энтропии текста  на английском языке | Энтропия разработанной программы для л.р.1 | Макс. значение энтропии разработанной программы для л.р.1 |
|  |  |  |  |

5. Проанализировать полученные результаты.