**Министерство образования и науки Российской Федерации**

федеральное государственное автономное образовательное учреждение высшего образования

**«САНКТ-ПЕТЕРБУРГСКИЙ НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ**

**УНИВЕРСИТЕТ ИНФОРМАЦИОННЫХ ТЕХНОЛОГИЙ,**

**МЕХАНИКИ И ОПТИКИ»**

Факультет среднего профессионального образования

**ОТЧЕТ**

**по Лабораторной работе №2**

**«Статистические алгоритмы сжатия данных без потерь»**

Специальность "Программирование в компьютерных системах" (09.02.03)

|  |  |
| --- | --- |
| Проверил:  Новиков В.В. \_\_\_\_\_\_\_\_\_\_  Кузнецова А.Л. \_\_\_\_\_\_\_\_\_\_  «\_\_\_\_\_»\_\_\_\_\_\_\_\_\_\_\_\_2018 г.  Оценка \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ | Выполнил:  Студент группы Y2235  Матвеенко Д.В.    \_\_\_\_\_\_\_\_\_\_\_  (подпись) |

Санкт-Петербург

2017/2018

ПОСТАНОВКА ЗАДАЧИ

**Цель лабораторной работы –** понять принципы сжатия данных путем применения статистических алгоритмов сжатия данных без потерь (на примере алгоритма Хаффмана).

**Задачи:**

1. Открыть из директории «10» текстовый файл в соответствии с вариантом.

2. Составить частотную таблицу, а затем кодовую таблицу по алгоритму Хаффмана.

3. Вычислить коэффициент сжатия данных алгоритмом.

4. Сравнить результаты сжатия с исходным текстом (объем данных), считая, что он закодирован в: а) ASCII; б) UTF-8.

ОТЧЕТ ПО ЛАБОРАТОРНОЙ РАБОТЕ №2

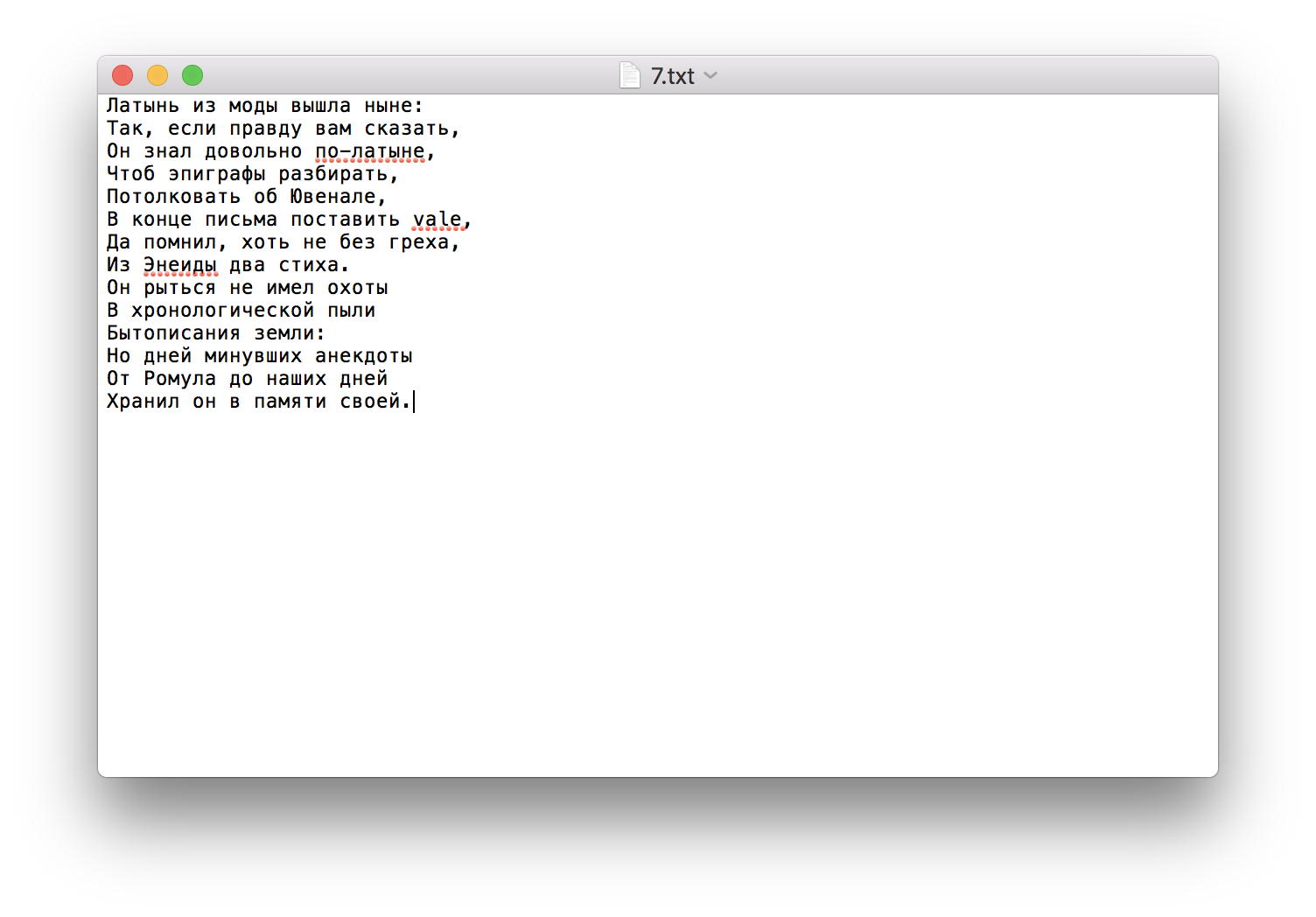
«Статистические алгоритмы сжатия данных без потерь»

**Вариант выполняемого задания:** 17 (7)

**Текст задания:**

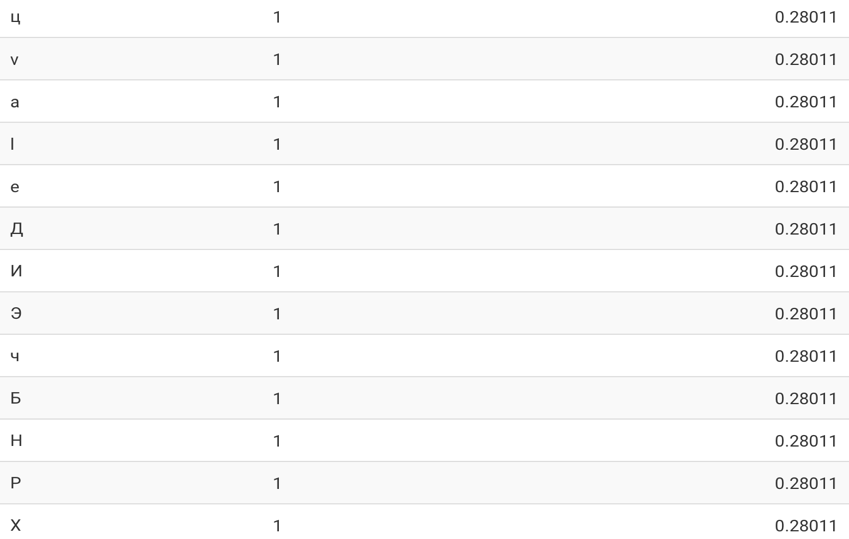
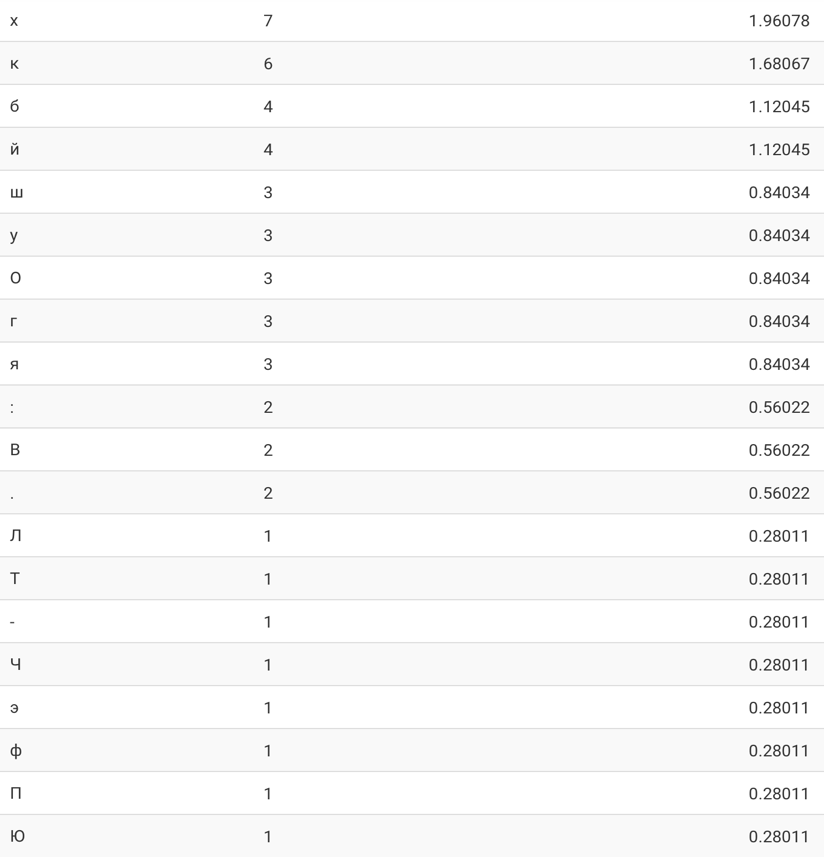
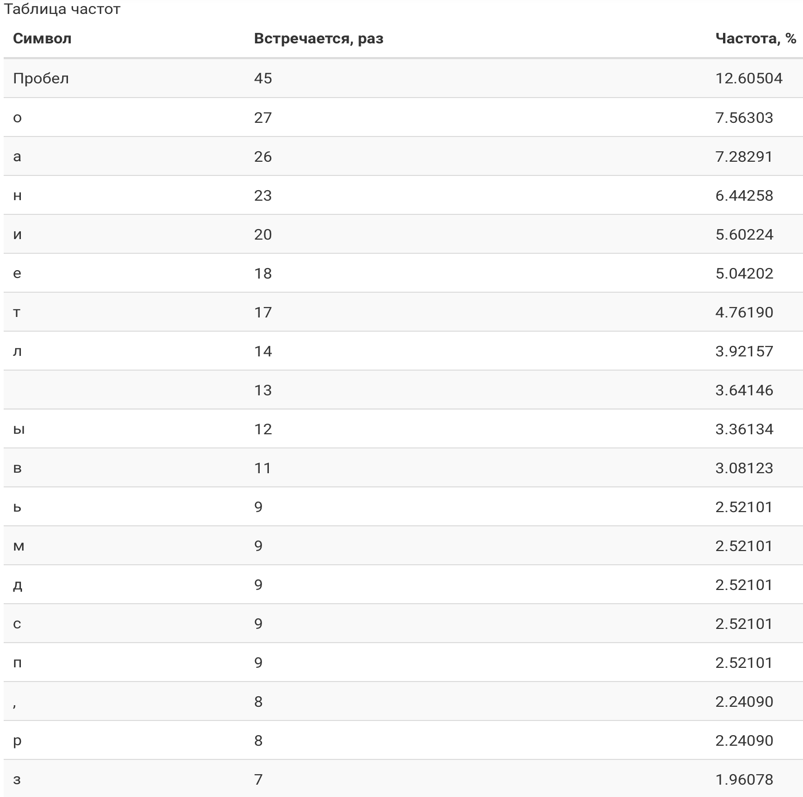
1. Открыть из директории «10» текстовый файл в соответствии с вариантом.
2. Составить частотную таблицу, а затем кодовую таблицу по алгоритму Хаффмана.
3. Вычислить коэффициент сжатия данных алгоритмом.
4. Сравнить результаты сжатия с исходным текстом (объем данных), считая, что он закодирован в: а) ASCII; б) UTF-8.

Ход работы

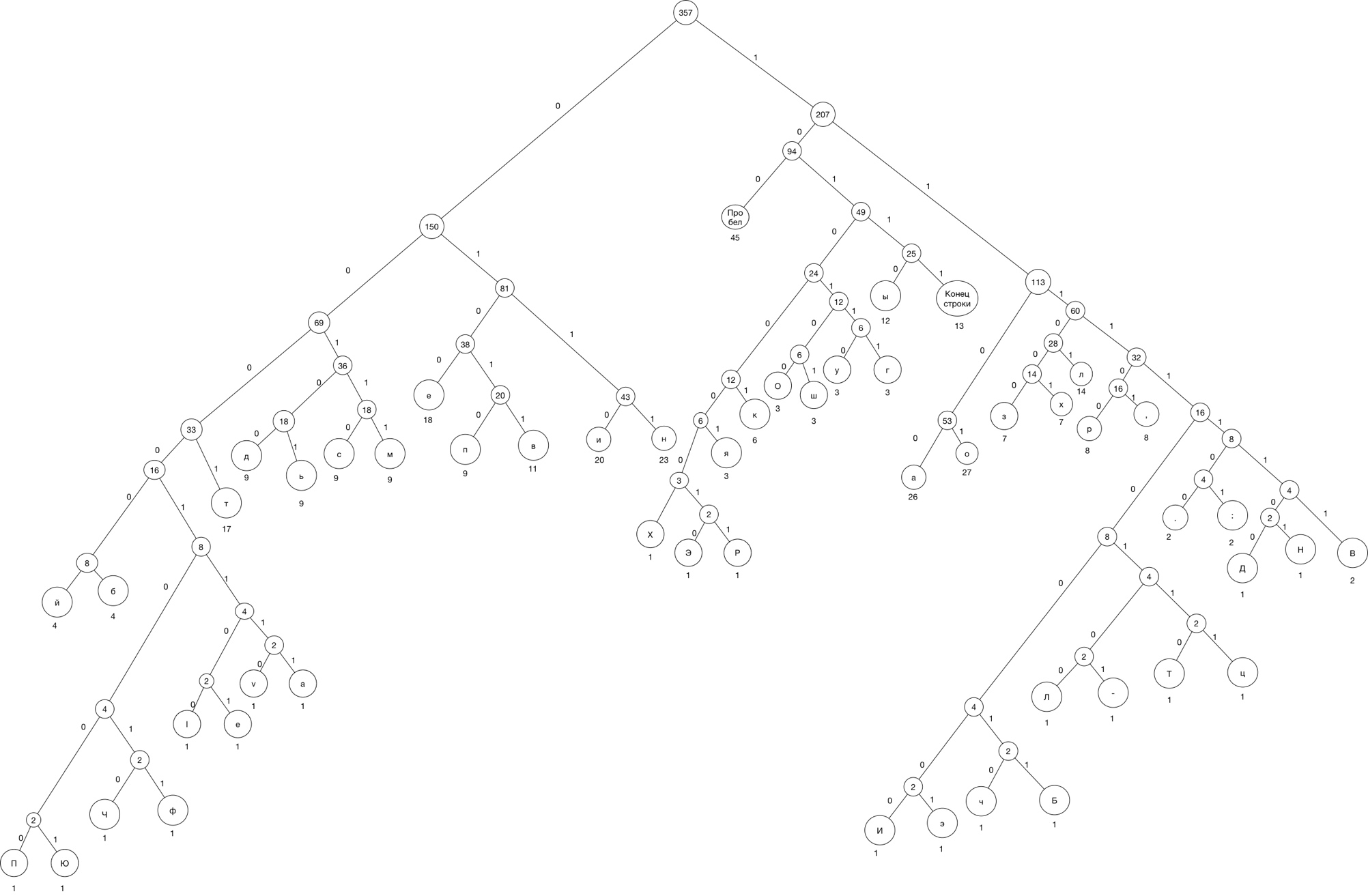
1. Для выполнения лабораторной был открыт текстовый файл под номером 7. Текст файла представлен ниже (Рисунок 1).

*Рисунок 1 – Текст файла*

1. Составим таблицу по убыванию частоты символов. Частотная таблица представлена ниже (Рисунок 2).

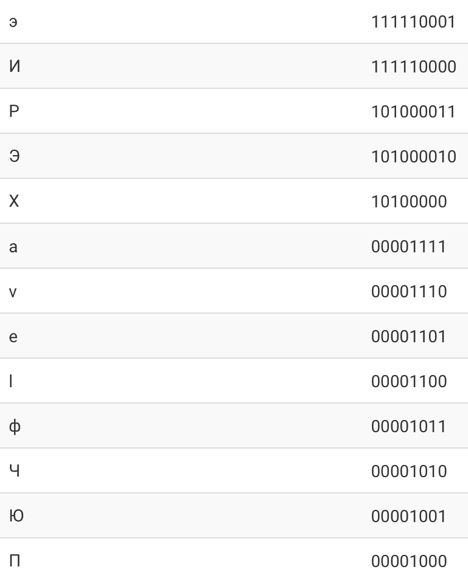
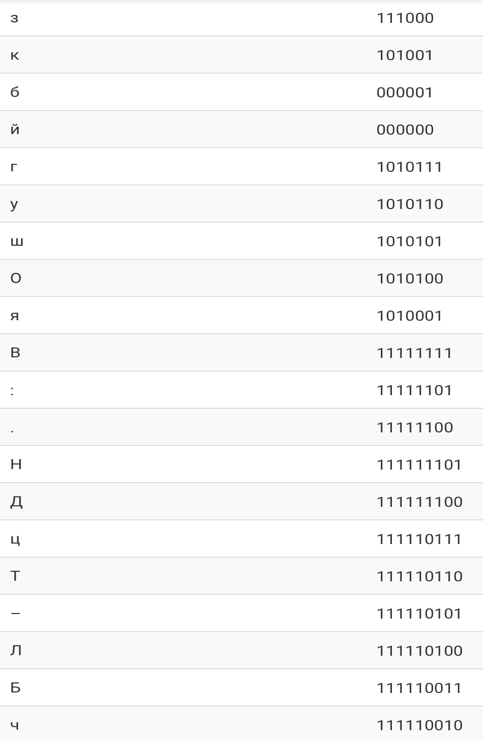
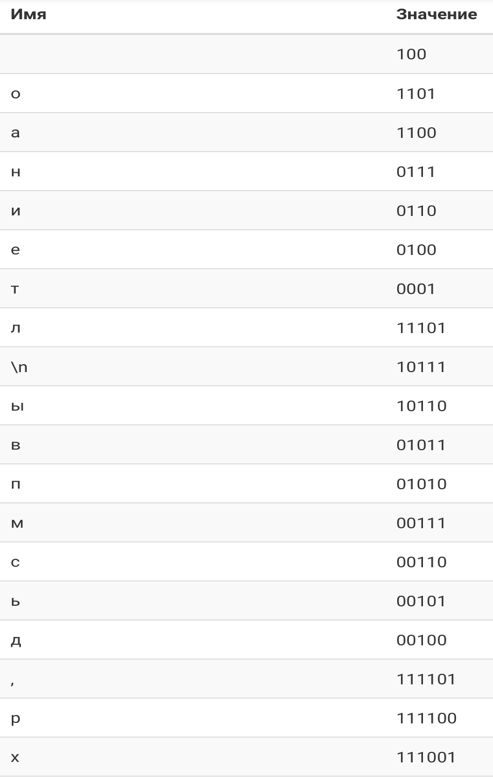
 *Рисунок 2 – Частотная таблица*

1. Построим дерево Хаффмана по частотной таблице. Запишем коды символов. Дерево Хаффмана представлено ниже (Рисунок 3).



*Рисунок 3 – Дерево Хаффмана*

4. В соответствии с деревом Хаффмана была составлена кодовая таблица символов (Рисунок 4).



*Рисунок 4 – Кодовая таблица*

5. Для нахождения коэффициентов сжатия были выполнены следующие действия:

1. Был найден размер исходного текста в кодировке ASCII - 2856 бит.
2. Был найден размер исходного текста в кодировке UTF-8 - 5120 бит.
3. Был найден размер закодированного текста - 1568 бит.

Коэффициенты сжатия текста равны:

1. ~1.82 (Кодировка текста – ASCII)
2. ~3.26 (Кодировка текста – UTF-8)

ЗАКЛЮЧЕНИЕ

В ходе работы были изучены статистические алгоритмы сжатия без потерь, приобретены практические навыки построения дерева Хаффмана, составления статистических таблиц.