# МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ

# ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ

# «БЕЛГОРОДСКИЙ ГОСУДАРСТВЕННЫЙ ТЕХНОЛОГИЧЕСКИЙ УНИВЕРСИТЕТ им. В. Г. ШУХОВА» (БГТУ им. В.Г. Шухова)

Кафедра программного обеспечения вычислительной техники и автоматизированных систем

# Лабораторная работа №3

по дисциплине: Теория информации тема: «Исследование возможности применения методов энтропийного кодирования для обработки двоичных последовательностей»

Выполнил: ст. группы ПВ-223 Дмитриев Андрей Александрович

Проверил:

Твердохлеб Виталий Викторович

# Содержание:

- 1. Тема лабораторной работы.
- 2. Цель лабораторной работы.
- 3. Текст задания к работе.
- 4. Выполнение задания.
  - а. Аналитика касательно построения кодов для исходной двоичной последовательности
  - b. Примеры кодовой реализации п.3, п.3, п.4 и п.6.
  - с. Результаты обработки кодов, полученных в п.5.
  - d. Текстовая последовательность, восстановленная к читаемомувиду.
  - е. Общие выводы.
- 5. Вывод по работе.

**Цель лабораторной работы:** изучить возможность применения методов энтропийного кодирования для обработки двоичных последовательностей. Написать и отладить программу составления кода для каждого символа методом Хаффмана и методом Шеннона-Фано, кодирования и декодирования двоичной последовательности. Сравнить время работы программы и коэффициенты сжатия при использовании метода Хаффманаи метода Шеннона-Фано.

# Текст задания к работе:

- 1. Открыть файл Лабораторная работа 3 (задание).txt. Рассмотреть возможность построения кода по методам Хаффмана и Шеннона-Фано для бинарной последовательности. Сделать выводы.
  - 2. Рассмотреть варианты обработки цепочек символов, а именно:
  - 2 символа;
  - 4 символа;
  - 8 символов.

Для этого разработать консольное приложение, разбивающее сплошной массив символов на цепочки заданной длины.

- 3. Рассматривая каждую цепочку (2, 4 и 8 символов длиной) как отдельный символ, построить коды по методу Хаффмана и Шеннона-Фано.
- 4. Составить последовательности из полученных кодов символов для каждого случая.
- 5. По результатам работы в п.3 сделать выводы по поводу полученных результатов для каждого из методов (простота, скорость, полученные результаты (рассчитать коэффициенты сжатия)).
- 6. Написать программу, восстанавливающую последовательности, полученные в п.3 в исходный вид согласно вариантам, приведенным в п.2.
- 7. Восстановить исходный текст из полученных последовательностей, пользуясь сервисом https://onlineutf8tools.com/convert-binary-to-utf8.

# Ход работы:

# **№**1

В файле содержится бинарная последовательность (0 и 1). Если строить код для каждого символа по отдельности, то каждому символу будет соответствовать код длиною 1 символ. Мы можем составлять код, взяв за символ последовательность нулей и единиц определенной длины. Если мы возьмём длину 2, то алфавит будет содержать максимум 4 символа, если длина будет равна 4, то алфавит будет состоять не больше, чем из 16 символов. Т. е. если длина последовательности из нулей и единиц, учитываемой как один символ, будет равна п, то алфавит сообщения будет состоять из 2<sup>n</sup> символов или меньше, если некоторые символы (последовательности нулей и единиц) не встретятся в сообщении.

# **№**2

# Код программы:

```
#include <iostream>
#include <fstream>
#include <vector>
#include <cmath>
int getNumberFromCharVector(const std::vector<char> &a)
    int i = a.size() - 1;
    int number = 0;
    for (auto &x : a)
        if (x == '1')
           number += pow(2, i);
        i--:
    return number;
std::vector<int> getSequencesOfNCharactersEach(const std::string &s, const int
n)
    std::vector<int> sequences;
    std::vector<char> a(n);
    int i = 1;
    for (auto &x : s)
        a[i] = x;
        if ((i + 1) \% n == 0)
            int p = getNumberFromCharVector(a);
            sequences.push back(p);
```

```
i = 0;
        else
            i++;
    return sequences;
void outputVector(const std::vector<int> &r)
    for (auto &x : r)
        std::cout << x << ' ';
    std::cout << '\n';</pre>
std::vector<int> getBinaryNumberNotation(std::vector<int> &a, int i, int n)
    int digit = n & 1;
    if (n == 0)
        return a;
    else
        getBinaryNumberNotation(a, i + 1, n >> 1);
        a[i] = digit;
    return a;
void outputVectorSequencesForReading(const std::vector<int> &r, const int
&length)
    for (auto &x : r)
        std::vector<int> a(length, 0);
        a = getBinaryNumberNotation(a, 0, x);
        for (auto &y : a)
            std::cout << y;</pre>
        std::cout << ' ';</pre>
    std::cout << '\n';</pre>
}
int main()
    std::string s;
    std::ifstream f0("T:\\2kurs2sem\\InformTheor\\lab3\\TI_3.txt");
    f0 >> s;
    f0.close();
```

```
int length = 2;
std::vector<int> r = getSequencesOfNCharactersEach(s, length);
std::cout << "In computer memory:\n";
outputVector(r);
std::cout << "\nFor reading:\n";
outputVectorSequencesForReading(r, length);
return 0;
}</pre>
```

Результат работы программы:

#### По 2 символа:

```
In computer memory:
1 2 2 0 1 0 2 1 1 2 2 0 1 1 2 2 3 2 2 0 3 0 0 1 1 2 2 0 1 1 2 2 3 2 2 0 3 0 0 0
0 1 0 0 1 2 2 0 3 0 0 0 3 2 2 0 1 1 2 1 1 2 2 0 1 1 3 0 1 2 2 0 3 0
                                                                0
3 0 0 1 1 2 2 0 1 1 2 2 3 2 2 0 1 1 3 1 2 1 1 2 0 1 0 0 1 2 2 0 1 1 2 1 1 2 2 0
   3 0 1 2 2 0 1 1 2 3 3 2 2 0 1 1 2 3 1 2 2 0 1 1 2 0 1 2 2 0 1 1 3 1 2 1
 100122011311220300012203013322030
 001122011223220113121001220113001001220
1 1 2 1 3 2 2 0 3 0 0 1 3 2 2 0 1 1 2 2 1 2 2 0 1 1 2 2 3 2 2 0 1 1 3 1 2 1 0 0
1 2 2 0 1 1 3 2 3 2 2 0 1 1 2 0 0 1 0 0 1 2 2 0 3 0 0 0 1 2 2 0 1 1 2 0 1 2 2 0
1 1 2 3 3 2 2 0 1 1 3 2 3 2 2 0 3 0 1 1 3 2 2 0 1 1 2 2 2 1 0 0 1 2 2 0 1 1 3 1
 2 2 0 1 1 2 0 1 2 2 0 1 1 2 2 1 2 2 0 3 0 1 1 2 1 1 3 0 1 0 0 1 2 2 0 1 1 0 1
1 2 2 0 1 1 3 3 0 1 0 0 1 2 2 0 1 1 2 3 1 2 2 0 1 1 2 0 1 2 2 0 1 1 3 1 3 2 2 0
1 1 3 3 1 2 2 0 1 1 2 0 3 2 2 0 1 1 3 2 3 2 2 0 3 0 1 1 3 2 2 0 1 1 3 2 0 1 0 0
   2 0 3 0 0 1 1 2 2 0 1 1 3 3 1 2 2 0 1 1 3 2 3 2 2 0 1 1 2 2 3 2 2 0 1 1 3 2
3 2 2 0 3 0 1 2 1 2 2 0 1 1 3 1 1 2 2 0 1 1 3 0 1 2 2 0 1 1 3 2 0 1 0 0 1 2 2 0
   2 1 3 2 2 0 1 1 3 3 1 2 2 0 1 1 3 1 3 2 2 0 1 1 3 3 1 2 2 0 3 0 0 0 3 2 2 0
 1 3 1 1 2 2 0 1 1 3 3 1 2 2 0 1 1 3 2 0 1 1 2 0 1 0 0 1 2 2 0 3 0 0
 1 3 3 0 1 0 0 1 2 2 0 1 1 2 1 3 2 2 0 3 0 0 0
                                             1 2 2 0 3 0 0 1 3
                                                              2
                                                                2 0 1 1
 2 2 0 3 0 1 1 3 2 2 0 1 1 3 2 0 1 0 0 1 2 2 0 1 1 2 0 3 2 2 0 1 1 2 0 1 2 2 0
 000322011331220112112203011322011320100
 2 2 0 3 0 0 0 1 2 2 0 1 1 2 0 1 2 2 0 3 0 0 0 3 2 2 0 1 1 3 1 1 2 2 0 1 1 2 0
1 2 2 0 3 0 0 1 1 2 2 0 1 1 3 3 1 2 2 0 1 1 3 2 0 1 0 0 1 2 2 0 3 0 0 0 1 2 2 0
1 1 2 0 1 2 2 0 3 0 0 0 3 2 2 0 1 1 3 3 3 2 2 0 1 1 2 2 3 2 2 0 1 1 2 1 1 2 2 0
1 1 2 0 1 2 2 0 1 1 3 1 2 1 0 0 1 2 2 0 1 1 3 3 1 2 2 0 1 1 3 2 2 1 0 0 1 2 2 0
3 0 0 0 3 2 2 0 1 1 2 1 1 2 2 0 1 1 3 3 1 2 2 0 3 0 1 3 0 1 0 0 1 2 2 0 1 1 2 0
3 2 2 0 1 1 3 3 1 2 2 0 1 1 2 2 3 2 2 0 1 1 2 1 1 2 2 0 3 0 0 1 3 2 2 0 3 0 1 3
 1 0 0 1 2 2 0 1 1 3 3 3 2 2 0 1 1 2 2 3 2 2 0 3 0 0 0 3 2 2 0 1 1 2 2 3 2 2 0
 1 3 2 3 2 2 0 1 1 3 1 1 2 2 0 3 0 0 1 2 1 1 3 0 1 0 0 1 2 2 0 1 1 0
                                                                  2 1 2 2 0
 1 3 3 1 2 2 0 1 1 3
                    2
                      3
                        2
                          2
                                1 2 0 1 2 2 0
                                             3 0 0 0 1 2 2 0 1
                            0
                              1
                                                              1
                                                                3
                                                                  0 0
1 2 2 0 3 0 0 3 3 2 2 0
                      3 0 0 1 3 2 2 0 3 0 0 1 1 2 2 0 3 0 1 2 0
     3 3 2 2 0 1 1 2 0
                      1
                        2
                          2 0 1 1 3 2 1 2 2 0
                                             1 1 2 2 3 2 2 0 3
                                                              0
                                                                0
                                                                  1
 1 3 2 3 2 2 0 1 1 3 3 0
                        100122011321220113012201121
 2 2 0 1 1 2 0 1 2 2 0 1 1 3 1 3 2 2 0 1 1 3 0 0 1 0 0 1 2 2 0 3
                                                              0
                                                                0
                                                                  0 3 2 2 0
1 1 3 1 1 2 2 0 1 1 2 1 1 2 2 0 1 1 3 3 1 2 2 0 1 1 2 3 3 2 2 0 3 0 1 2 0 1 0 0
 2 2 0 1 1 3 3 1 2 2 0 1 1 2 1 3 2 2 0 3 0 0 0 1 2 2 0 1 1 3 3 1 2 2 0 1 1 3 2
 2 2 0 1 1 3 2 3 2 2 0 3 0 1 1 3 2 2 0 1 1 2 2 2 1 0 0 1 2 2 0 1 1 2 0 3 2 2 0
 1 2 2 3 2 2 0 1 1 3 1 3 2 2 0 3 0 1 1 3 2 2 0 1 1 2 2 2 1 0 0 1 2 2 0 3 0 0 2
```

3 2 2 0 1 1 3 1 3 2 2 0 1 1 3 3 1 2 2 0 1 1 3 3 3 2 2 0 3 0 1 2 1 2 2 0 3 0 1 3 2 1 0 0 1 2 2 0 3 0 0 0 3 2 2 0 1 1 3 2 3 2 2 0 1 1 2 2 3 2 2 0 1 1 2 1 3 2 2 0 1 1 2 0 0 1 1 2 0 1 0 0 1 2 2 0 1 1 3 3 1 2 2 0 1 1 2 0 3 2 2 0 1 1 3 0 1 2 2 0 1 1 3 1 3 2 2 0 3 0 1 2 1 2 2 0 1 1 3 2 3 2 2 0 1 1 3 3 0 1 0 0 1 2 2 0 3 0 0 0 2 2 0 3 0 1 1 3 2 2 0 1 1 3 3 3 2 2 0 1 1 2 0 1 2 2 0 1 1 2 1 1 2 2 0 3 0 1 0 1 2 2 0 1 1 3 0 1 2 2 0 1 1 2 2 3 2 2 0 3 0 0 0 3 2 2 0 3 0 1 3 2 1 0 0 1 2 2 0 1 1 3 2 3 2 2 0 1 1 2 0 0 1 0 0 1 2 2 0 3 0 0 1 1 2 2 0 3 0 0 0 1 2 2 0 1 1 3 3 1 2 2 0 3 0 0 1 1 2 2 0 3 0 0 1 3 2 2 0 1 1 2 0 1 2 2 0 3 0 0 0 1 2 2 0 3 0 1 1 2 1 1 2 0 1 0 0 1 2 2 0 1 1 3 2 3 2 2 0 1 1 2 0 0 1 0 0 1 2 2 0 3 0 0 1 3 2 2 0 1 1 3 1 3 2 2 0 1 1 3 0 1 2 2 0 3 0 0 3 1 2 2 0 3 0 0 1 2 1 1 2 0 1 0 0 1 2 2 0 1 1 3 2 3 2 2 0 1 1 2 0 0 1 0 0 1 2 2 0 3 0 1 2 3 2 2 0 1 1 3 1 1 2 2 0 1 1 3 0 1 2 2 0 1 1 3 3 3 2 2 0 1 1 2 0 1 2 2 0 1 1 2 3 1 2 2 0 1 1 3 0 0 1 1 2 0 1 0 0 1 2 2 0 1 1 3 1 3 2 2 0 1 1 3 3 1 2 2 0 3 0 1 0 1 2 2 0 1 1 2 0 1 2 2 0 1 1 2 2 1 2 2 0 1 1 2 2 3 2 2 0 1 1 3 0 2 1 0 0 1 2 2 0 1 1 3 0 0 1 0 0 1 2 2 0 1 1 3 3 3 2 2 0 3 0 0 0 1 2 2 0 1 1 3 3 1 2 2 0 3 0 0 2 3 2 2 0 1 1 3 3 1 2 2 0 1 1 2 3 1 2 2 0 1 1 3 0 1 2 2 0 3 0 0 2 2 1 1 3

#### For reading:

```
10
11 00 10 10
<u>10 01 01 00 10 10 0</u>1 00 10 01 01 00 10 10 01 01 10 01 01 0<del>0</del> 10 10 01 01 11 01
   01
11 00 00 01 01 10 10
```

## По 4 символа:

#### In computer memory:

6 8 4 9 6 8 5 10 14 8 12 1 6 8 5 10 14 8 12 0 1 0 6 8 12 0 14 8 5 9 6 8 5 12 6 8 12 0 14 8 12 1 6 8 5 10 14 8 5 13 9 6 1 0 6 8 5 9 6 8 5 12 6 8 5 11 14 8 5 11 6 8 5 8 6 8 5 13 9 6 1 0 6 8 5 13 6 8 12 0 6 8 12 7 14 8 12 2 14 8 12 1 6 8 5 10 14 8 5 10 14 8 5 13 9 0 6 8 5 12 1 0 6 8 5 9 14 8 12 1 14 8 5 10 6 8 5 10 14 8 5 13 9 0 6 8 5 14 14 8 5 8 1 0 6 8 12 0 6 8 5 8 6 8 5 11 14 8 5 14 14 8 12 5 14 8 5 10 9 0 6 8 5 13 14 8 5 8 6 8 5 10 6 8 5 10 6 8 5 11 6 8 5

8 6 8 5 13 14 8 5 15 6 8 5 8 14 8 5 14 14 8 12 5 14 8 5 14 1 0 6 8 12 1 6 8 5 15 6 8 5 14 14 8 5 10 14 8 5 14 14 8 12 6 6 8 5 13 6 8 5 12 6 8 5 14 1 0 6 8 5 9 14 8 5 15 6 8 5 13 14 8 5 15 6 8 12 0 14 8 5 13 6 8 5 15 6 8 5 14 1 6 1 0 6 8 12 1 6 8 5 15 1 0 6 8 5 9 14 8 12 0 6 8 12 1 14 8 5 8 14 8 12 5 14 8 5 14 1 0 6 8 5 8 14 8 5 8 6 8 12 0 14 8 5 15 6 8 5 9 6 8 12 5 14 8 5 14 1 0 6 8 12 0 6 8 5 8 6 8 12 0 14 8 5 13 6 8 5 8 6 8 12 1 6 8 5 15 6 8 5 14 1 0 6 8 12 0 6 8 5 8 6 8 12 0 <u>14 8 5 15 14 8 5 10 14 8 5 9 6 8 5 8 6 8 5 13 9 0 6 8 5 15 6 8 5 14 9 0 6 8 12 0</u> 14 8 5 9 6 8 5 15 6 8 12 7 1 0 6 8 5 8 14 8 5 15 6 8 5 10 14 8 5 9 6 8 12 1 14 8 12 7 1 0 6 8 5 15 14 8 5 10 14 8 12 0 14 8 5 10 14 8 5 14 14 8 5 13 6 8 12 1 9 7 1 0 6 8 5 2 6 8 5 15 6 8 5 14 14 8 5 8 6 8 12 0 6 8 5 12 1 0 6 8 12 3 14 8 12 1 14 8 12 1 6 8 12 6 1 0 6 8 5 11 14 8 5 8 6 8 5 14 6 8 5 10 14 8 12 1 6 8 5 14 14 8 5 15 1 0 6 8 5 14 6 8 5 12 6 8 5 9 14 8 5 8 6 8 5 13 14 8 5 12 1 0 6 8 12 0 14 8 5 13 6 8 5 9 6 8 5 15 6 8 5 11 14 8 12 6 1 0 6 8 5 15 6 8 5 9 14 8 12 0 6 8 5 15 6 8 5 14 6 8 5 14 14 8 12 5 14 8 5 10 9 0 6 8 5 8 14 8 5 10 14 8 5 13 14 8 12 5 14 8 5 10 9 0 6 8 12 2 14 8 5 13 14 8 5 15 6 8 5 15 14 8 12 6 6 8 12 7 9 0 6 8 12 0 14 8 5 14 14 8 5 10 14 8 5 9 14 8 5 8 1 6 1 0 6 8 5 15 6 8 5 8 14 8 5 12 6 8 5 13 14 8 12 6 6 8 5 14 14 8 5 15 1 0 6 8 12 0 14 8 12 5 14 8 5 15 14 8 5 8 6 8 5 9 6 8 12 4 6 8 5 12 6 8 5 10 14 8 12 0 14 8 12 7 9 0 6 8 5 14 14 8 5 8 1 0 6 8 12 1 6 8 12 0 6 8 5 15 6 8 12 1 6 8 12 1 14 8 5 8 6 8 12 0 6 8 12 5 9 6 1 0 6 8 5 14 14 8 5 8 1 0 6 8 12 1 14 8 5 13 14 8 5 12 6 8 12 3 6 8 12 1 9 6 1 0 6 8 5 14 14 8 5 8 1 0 6 8 12 6 14 8 5 13 6 8 5 12 6 8 5 15 14 8 5 8 6 8 5 11 6 8 5 12 1 6 1 0 6 8 5 13 14 8 5 15 6 8 12 4 6 8 5 8 6 8 5 10 6 8 5 10 14 8 5 12 9 0 6 8 5 12 1 0 6 8 5 15 14 8 12 0 6 8 5 15 6 8 12 2 14 8 5 15 6 8 5 11 6 8 5 12 6 8 12 2 9 7

## For reading:

0110 0001 0010 1001 0110 0001 1010 0101 0111 0001 0011 1000 0110 0001 1010 0101 0111 0001 0011 0000 1000 0000 0110 0001 0011 0000 0111 0001 1010 1001 0110 0001 1010 0011 0110 0001 0011 0000 0111 0001 0011 1000 0110 0001 1010 0101 0111 0001 1010 1011 1001 0110 1000 0000 0110 0001 1010 1001 0110 0001 1010 0011 0110 0001 1010 1101 0111 0001 1010 1101 0110 0001 1010 0001 0110 0001 1010 1011 1001 0110 1000 0000 0110 0001 1010 1011 0110 0001 0011 0000 0110 0001 0011 1110 0111 0001 0011 0100 0111 0001 0011 1000 0110 0001 1010 0101 0111 0001 1010 1011 1001 0000 0110 0001 1010 0011 1000 0000 0110 0001 1010 1001 0111 0001 0011 1000 0111 0001 1010 0101 0110 0001 1010 0101 0111 0001 1010 1011 1001 0000 0110 0001 1010 0111 0111 0001 1010 0001 1000 0000 0110 0001 0011 0000 0110 0001 1010 0001 0110 0001 1010 1101 0111 0001 1010 0111 0111 0001 0011 1010 0111 0001 1010 0101 1001 0000 0110 0001 1010 1011 0111 0001 1010 0001 0110 0001 1010 0101 0110 0001 0011 1010 1001 1110 1000 0000 0110 0001 1010 1000 0110 0001 1010 1111 1000 0000 0110 0001 1010 1101 0110 0001 1010 0001 0110 0001 1010 1011 0111 0001 1010 1111 0110 0001 1010 0001 0111 0001 1010 0111 0111 0001 0011 1010 0111 0001 1010 0111 1000 0000 0110 0001 0011 1000 0110 0001 1010 1111 0110 0001 1010 0111 0111 0001 1010 0101 0111 0001 1010 0111 0111 0001 0011 0110 0110 0001 1010 1011 0110 0001 1010 0011 0110 0001 1010 0111 1000 0000 0110 0001 1010 1001 0111 0001 1010 1111 0110 0001 1010 1011 0111 0001 1010 1111 0110 0001 0011 0000 0111 0001 1010 1011 0110 0001 1010 1111 0110 0001 1010 0111 1000 0110 1000 0000 0110 0001 0011 1000 0110 0001 1010 1111 1000 0000 0110 0001 1010 1001 0111 0001 0011 0000 0110 0001 0011 1000 0111 0001 1010 0001 0111 0001 0011 1010 0111 0001 1010 0111 1000 0000 0110 0001 1010 0001 0111 0001 1010 0001 0110 0001 0011 0000 0111 0001 1010 1111 0110 0001 1010 1001 0110 0001 0011 1010 0111 0001 1010 0111 1000 0000 0110 0001 0011 0000 0110 0001 1010 0001 0110 0001 0011 0000 0111 0001 1010 1011 0110 0001 1010 0001 0110 0001 0011 1000 0110 0001 1010 1111 0110 0001 1010 0111 1000 0000 0110 0001

```
0011 0000 0110 0001 1010 0001 0110 0001 0011 0000 0111 0001 1010 1111 0111 0001
1010 0101 0111 0001 1010 1001 0110 0001 1010 0001 0110 0001 1010 1011 1001 0000
0110 0001 1010 1111 0110 0001 1010 0111 1001 0000 0110 0001 0011 0000 0111 0001
1010 1001 0110 0001 1010 1111 0110 0001 0011 1110 1000 0000 0110 0001 1010 0001
0111 0001 1010 1111 0110 0001 1010 0101 0111 0001 1010 1001 0110 0001 0011 1000
0111 0001 0011 1110 1000 0000 0110 0001 1010 1111 0111 0001 1010 0101 0111 0001
0011 0000 0111 0001 1010 0101 0111 0001 1010 0111 0111 0001 <u>1010 1011 0110 0001</u>
0011 1000 1001 1110 1000 0000 0110 0001 1010 0100 0110 0001 1010 1111 0110 0001
1010 0111 0111 0001 1010 0001 0110 0001 0011 0000 0110 0001 1010 0011 1000 0000
0110 0001 0011 1100 0111 0001 0011 1000 0111 0001 0011 1000 0110 0001 0011 0110
1000 0000 0110 0001 1010 1101 0111 0001 1010 0001 0110 0001 1010 0111 0110 0001
1010 0101 0111 0001 0011 1000 0110 0001 1010 0111 0111 0001 1010 1111 1000 0000
0110 0001 1010 0111 0110 0001 1010 0011 0110 0001 1010 1001 0111 0001 1010 0001
0110 0001 1010 1011 0111 0001 1010 0011 1000 0000 0110 0001 0011 0000 0111 0001
1010 1011 0110 0001 1010 1001 0110 0001 1010 1111 0110 0001 1010 1101 0111 0001
0011 0110 1000 0000 0110 0001 1010 1111 0110 0001 1010 1001 0111 0001 0011 0000
0110 0001 1010 1111 0110 0001 1010 0111 0110 0001 1010 0111 0111 0001 0011 1010
0111 0001 1010 0101 1001 0000 0110 0001 1010 0001 0111 0001 1010 0101 0111 0001
1010 1011 0111 0001 0011 1010 0111 0001 1010 0101 1001 0000 0110 0001 0011 0100
0111 0001 1010 1011 0111 0001 1010 1111 0110 0001 1010 1111 0111 0001 0011 0110
0110 0001 0011 1110 1001 0000 0110 0001 0011 0000 0111 0001 1010 0111 0111 0001
1010 0101 0111 0001 1010 1001 0111 0001 1010 0001 1000 0110 1000 0000 0110 0001
1010 1111 0110 0001 1010 0001 0111 0001 1010 0011 0110 0001 1010 1011 0111 0001
0011 0110 0110 0001 1010 0111 0111 0001 1010 1111 1000 0000 0110 0001 0011 0000
0111 0001 0011 1010 0111 0001 1010 1111 0111 0001 1010 0001 0110 0001 1010 1001
0110 0001 0011 0010 0110 0001 1010 0011 0110 0001 1010 0101 0111 0001 0011 0000
0111 0001 0011 1110 1001 0000 0110 0001 1010 0111 0111 0001 1010 0001 1000 0000
0110 0001 0011 1000 0110 0001 0011 0000 0110 0001 1010 1111 0110 0001 0011 1000
0110 0001 0011 1000 0111 0001 1010 0001 0110 0001 0011 0000 0110 0001 0011 1010
1001 0110 1000 0000 0110 0001 1010 0111 0111 0001 1010 0001 1000 0000 0110 0001
0011 1000 0111 0001 1010 1011 0111 0001 1010 0011 0110 0001 0011 1100 0110 0001
0011 1000 1001 0110 1000 0000 0110 0001 1010 0111 0111 0001 1010 0001 1000 0000
0110 0001 0011 0110 0111 0001 1010 1011 0110 0001 1010 0011 0110 0001 1010 1111
0111 0001 1010 0001 0110 0001 1010 1101 0110 0001 1010 0011 1000 0110 1000 0000
0110 0001 1010 1011 0111 0001 1010 1111 0110 0001 0011 0010 0110 0001 1010 0001
0110 0001 1010 0101 0110 0001 1010 0101 0111 0001 1010 0011 1001 0000 0110 0001
1010 0011 1000 0000 0110 0001 1010 1111 0111 0001 0011 0000 0110 0001 1010 1111
0110 0001 0011 0100 0111 0001 1010 1111 0110 0001 1010 1101 0110 0001 1010 0011
0110 0001 0011 0100 1001 1110
```

#### По 8 символов:

In computer memory:

```
104 73 104 90 232 193 104 90 232 192 16 104 192 232 89 104 92 104 192 232 193 104 90 232 93 150 16 104 89 104 92 104 91 232 91 104 88 104 93 150 16 104 93 104 192 104 199 232 194 232 193 104 90 232 93 144 104 92 16 104 89 232 193 232 90 104 90 232 93 144 104 94 232 88 16 104 192 104 88 104 91 232 94 232 197 232 90 144 104 93 232 88 104 90 104 197 151 16 104 81 104 95 16 104 91 104 88 104 93 232 95 104 88 232 94 232 197 232 94 16 104 193 104 95 104 94 232 90 232 94 232 198 104 93 104 92 104 94 16 104 89 232 95 104 193 232 88 232 197 232 94 16 104 94 22 16 104 193 104 95 16 104 89 232 192 104 193 232 88 232 197 232 94 16
```

104 88 232 88 104 192 232 95 104 89 104 197 232 94 16 104 192 104 88 104 192 232 93 104 88 104 193 104 95 104 94 16 104 192 104 88 104 192 232 95 232 90 232 89 104 88 104 93 144 104 95 104 94 144 104 192 232 89 104 95 104 199 16 104 88 232 95 104 90 232 89 104 193 232 199 16 104 95 232 90 232 192 232 90 232 94 232 93 104 193 151 16 104 82 104 95 104 94 232 88 104 192 104 92 16 104 195 232 193 232 193 104 198 16 104 91 232 88 104 94 104 90 232 193 104 94 232 95 16 104 94 104 92 104 89 232 88 104 93 232 92 16 104 192 232 93 104 89 104 95 104 91 232 198 16 104 95 104 89 232 192 104 95 104 94 104 94 232 197 232 90 144 104 88 232 90 232 93 232 197 232 90 144 104 194 232 93 232 95 104 95 232 198 104 199 144 104 192 232 94 232 90 232 89 232 88 22 16 104 95 104 88 232 92 104 93 232 198 104 94 232 95 16 104 192 232 197 232 95 232 88 22 16 104 95 104 88 232 92 104 93 232 198 104 94 232 95 16 104 192 232 197 232 95 232 88 104 89 104 196 104 92 104 90 232 192 232 199 144 104 94 232 88 16 104 193 104 192 104 95 104 193 104 193 232 88 104 192 104 197 150 16 104 94 232 88 16 104 193 232 93 232 92 104 195 104 193 150 16 104 94 232 88 16 104 198 232 93 104 92 104 95 232 88 104 91 104 92 22 16 104 95 104 194 232 95 104 196 104 88 104 90 104 90 232 92 144 104 92 16 104 95 232 192 104 95 104 194 232 95 104 191 104 92 104 194 151

#### For reading:

00111010 00001000 00010110 00000011 00010111 10111010 00010110 10011010 00010110

```
10011010 00010111 00000011 00010110 11111010 00010110 01111010 00010110 01111010
00010111 10111010 00010111 11111010 00010110 11111010 00010111 01100011 00010110
10011010 00010111 00011010 01101000 00001000 00010110 11111010 00010110 00011010
00010111 00111010 00010110 10111010 00010111 01100011 00010110 01111010 00010111
11111010 00001000 00010110 00000011 00010111 10100011 00010111 11111010 00010111
00001000 00010110 01111010 00010111 00011010 00001000 00010110 10000011 00010111
10111010 00010111 00111010 00010110 11000011 00010110 10000011 01101001 00001000
00010110 01111010 00010111 00011010 00001000 00010110 01100011 00010111 10111010
00111010 01101000 00001000 00010110 10111010 00010111 11111010 00010110 00100011
01000011 11101001
```

#### **№**3

# 1) Метод Хаффмана:

Код программы:

```
// возвращает отсортированный по неувозрастанию вектор содержащий
// структуру character, полученную после обработки вектора table,
// содержащего двоичные последовательности, записанные целыми числами
// поле symbol - двоичная последовательность, записанная целыми числами
std::vector<character> getTable(const std::vector<int> &table)
    std::vector<character> res;
    for (auto &x : table)
        int pos = getSymbolPosition(res, x);
        if (pos == -1)
            res.push_back(character(std::vector<int>{x}, 1,
std::vector<int>()));
        else
            res[pos].numbers++;
    std::sort(res.begin(), res.end(), comp);
    return res;
void reverseVector(std::vector<int> &v)
    for (int i = 0; i < v.size() / 2; i++)
        bool c = v[i];
        v[i] = v[v.size() - 1 - i];
        v[v.size() - 1 - i] = c;
    }
// метод Хаффмана
// возвращает таблицу, содержащую символ, его количество повторений, код
std::vector<character> theHuffmanMethod(const std::vector<int> &table)
    std::vector<character> res = getTable(table);
    std::vector<character> p = res;
    while (p.size() > 1)
        int n = p.size() - 1;
        for (int i = 0; i < p[n].symbol.size(); i++)
            int k = getSymbolPosition(res, p[n].symbol[i]);
            res[k].code.push_back(0);
        for (int i = 0; i < p[n].symbol.size(); i++)
```

```
{
            res[getSymbolPosition(res, p[n].symbol[i])].code.push_back(1);
        p[n].numbers += p[n + 1].numbers;
        for (int i = 0; i < p[n + 1].symbol.size(); i++)
            p[n].symbol.push_back(p[n + 1].symbol[i]);
        p.erase(p.cend());
        std::sort(p.begin(), p.end(), comp);
    for (auto &x : res)
        reverseVector(x.code);
    return res;
void outputSymbolCodes(const std::vector<character> &a, int codeLength)
    for (character element : a)
        std::cout << "Symbol: " << element.symbol[0] << " / ";</pre>
        std::vector<int> a(codeLength, 0);
        a = getBinaryNumberNotation(a, 0, element.symbol[0]);
        for (auto &x : a)
            std::cout << x;</pre>
        std::cout << "\nCode: ";</pre>
        for (int x : element.code)
            std::cout << x;</pre>
        std::cout << "\n\n";</pre>
    }
int main()
    std::string s;
    std::ifstream f0("T:\\2kurs2sem\\InformTheor\\lab3\\TI_3.txt");
   f0 >> s;
   f0.close();
    int length = 8;
    std::vector<int> r = getSequencesOfNCharactersEach(s, length);
    std::vector<character> res = theHuffmanMethod(r);
    outputSymbolCodes(res, length);
    return 0;
```

# По 2 символа:

Symbol: 2 / 01

Code: 11

Symbol: 1 / 10

Code: 10

Symbol: 0 / 00

Code: 01

Symbol: 3 / 11

Code: 00

# По 4 символа:

Symbol: 8 / 0001

Code: 10

Symbol: 5 / 1010

Code: 111

Symbol: 6 / 0110

Code: 110

Symbol: 14 / 0111

Code: 010

Symbol: 12 / 0011

Code: 000

Symbol: 0 / 0000

Code: 0110

Symbol: 1 / 1000

Code: 0010

Symbol: 9 / 1001

Code: 01110

Symbol: 15 / 1111

Code: 00111

Symbol: 10 / 0101

Code: 011111

Symbol: 13 / 1011

Code: 011110

Symbol: 11 / 1101

Code: 0011011

Symbol: 7 / 1110

Code: 0011010

Symbol: 2 / 0100

Code: 0011001

Symbol: 4 / 0010

Code: 00110001

Symbol: 3 / 1100 Code: 00110000

#### По 8 символов:

Symbol: 104 / 00010110 Code: 10 Symbol: 232 / 00010111 Code: 111 Symbol: 16 / 00001000 Code: 0111 Symbol: 95 / 11111010 Code: 0110 Symbol: 88 / 00011010 Code: 0100 Symbol: 94 / 01111010 Code: 0011 Symbol: 192 / 00000011 Code: 0001 Symbol: 90 / 01011010 Code: 11011 Symbol: 93 / 10111010 Code: 0000 Symbol: 193 / 10000011 Code: 11001 Symbol: 92 / 00111010 Code: 01011 Symbol: 89 / 10011010 Code: 01010 Symbol: 144 / 00001001 Code: 110101 Symbol: 197 / 10100011 Code: 110100 Symbol: 91 / 11011010 Code: 110001 Symbol: 198 / 01100011 Code: 001010

Code: 001000

Symbol: 199 / 11100011

```
Symbol: 194 / 01000011
Code: 1100001
Symbol: 150 / 01101001
Code: 1100000
Symbol: 151 / 11101001
Code: 0010110
Symbol: 22 / 01101000
Code: 0010011
Symbol: 196 / 00100011
Code: 00101111
Symbol: 195 / 11000011
Code: 00101110
Symbol: 82 / 01001010
Code: 00100100
Symbol: 81 / 10001010
Code: 001001011
Symbol: 73 / 10010010
Code: 0010010
```

# 2) Метод Шеннона-Фано Код программы:

```
#include "3_2.h"
std::vector<character> theShannonFanoMethod_(std::vector<character> a, int from,
int to)
    if (to - from > 1)
        int mid;
        if (to - from == 2)
            mid = from + 1;
        else
            int sum = 0;
            for (int i = from; i < to; ++i)
                sum += a[i].numbers;
            int halfOfSum = sum / 2;
            sum = 0;
            int i = from;
            while (sum < halfOfSum && i < to)</pre>
                sum += a[i].numbers;
```

```
i++;
            int k1 = abs(sum - a[i].numbers - halfOfSum);
            int k2 = abs(sum - halfOfSum);
            if (k1 < k2)
                mid = i;
            else
                mid = i + 1;
        for (int i = from; i < mid; i++)</pre>
            a[i].code.push_back(0);
        a = theShannonFanoMethod_(a, from, mid);
        for (int i = mid; i < to; i++)
            a[i].code.push back(1);
        a = theShannonFanoMethod_(a, mid, to);
    return a;
// возвращает таблицу, содержащую символ, его количество повторений, код
std::vector<character> theShannonFanoMethod(const std::vector<int> &table)
    std::vector<character> res = getTable(table);
    res = theShannonFanoMethod_(res, 0, res.size());
    return res;
int main()
    std::string s;
    std::ifstream f0("T:\\2kurs2sem\\InformTheor\\lab3\\TI_3.txt");
    f0 >> s;
    f0.close();
    int length = 2;
    std::vector<int> r = getSequencesOfNCharactersEach(s, length);
    std::vector<character> res = theShannonFanoMethod(r);
    outputSymbolCodes(res, length);
    return 0;
```

Результат работы программы:

По 2 символа:

```
Symbol: 2 / 01
Code: 000
Symbol: 1 / 10
Code: 001
Symbol: 0 / 00
Code: 01
Symbol: 3 / 11
Code: 1
```

# По 4 символа:

Symbol: 7 / 1110 Code: 111101

```
Symbol: 8 / 0001
Code: 000
Symbol: 5 / 1010
Code: 001
Symbol: 6 / 0110
Code: 01
Symbol: 14 / 0111
Code: 1000
Symbol: 12 / 0011
Code: 1001
Symbol: 0 / 0000
Code: 101
Symbol: 1 / 1000
Code: 11000
Symbol: 9 / 1001
Code: 11001
Symbol: 15 / 1111
Code: 1101
Symbol: 10 / 0101
Code: 11100
Symbol: 13 / 1011
Code: 11101
Symbol: 11 / 1101
Code: 111100
```

Symbol: 2 / 0100 Code: 1111100

Symbol: 4 / 0010 Code: 1111101

Symbol: 3 / 1100 Code: 111111

# По 8 символов:

Symbol: 104 / 00010110

Code: 000

Symbol: 232 / 00010111

Code: 001

Symbol: 16 / 00001000

Code: 010

Symbol: 95 / 11111010

Code: 011

Symbol: 88 / 00011010

Code: 100000

Symbol: 94 / 01111010

Code: 100001

Symbol: 192 / 00000011

Code: 10001

Symbol: 90 / 01011010

Code: 1001

Symbol: 93 / 10111010

Code: 1010

Symbol: 193 / 10000011

Code: 1011

Symbol: 92 / 00111010

Code: 110000

Symbol: 89 / 10011010

Code: 110001

Symbol: 144 / 00001001

Code: 11001

Symbol: 197 / 10100011

Code: 11010

Symbol: 91 / 11011010

Code: 11011

Symbol: 198 / 01100011

Code: 1110000

Symbol: 199 / 11100011

Code: 1110001

Symbol: 194 / 01000011

Code: 111001

Symbol: 150 / 01101001

Code: 11101

Symbol: 151 / 11101001

Code: 1111000

Symbol: 22 / 01101000

Code: 1111001

Symbol: 196 / 00100011

Code: 111101

Symbol: 195 / 11000011

Code: 1111100

Symbol: 82 / 01001010

Code: 1111101

Symbol: 81 / 10001010

Code: 1111110

Symbol: 73 / 10010010

Code: 111111

```
#include "4.h"
std::string
replaceCharactersWithTheirCodes(const std::vector<character> &table, const
std::vector<int> &s)
    std::vector<char> a;
    std::string res;
    for (auto &x : s)
        int pos = getSymbolPosition(table, x);
        if (pos >= 0)
            for (auto &y : table[pos].code)
                res.push_back('0' + y);
    return res;
int main()
    std::string s;
    std::ifstream f0("T:\\2kurs2sem\\InformTheor\\lab3\\TI_3.txt");
    f0 >> s;
    f0.close();
    int length = 2;
    std::vector<int> r = getSequencesOfNCharactersEach(s, length);
    std::vector<character> res = theHuffmanMethod(r);
    std::string sCode = replaceCharactersWithTheirCodes(res, r);
    std::cout << sCode;</pre>
    return 0;
```

Результат работы программы:

Метод Хаффмана

# По 2 символа:

# По 4 символа:

# По 8 символов:

## 1) Метод Шеннона-Фано

#### По 2 символа:

#### По 4 символа:

# По 8 символов:

## №5

## Код программы:

```
#include "5.h"

double getCompressionRatio(std::string s, const std::vector<character> &table)
{
```

```
int numberOfSymbols = s.size();
    int b = numberOfSymbols;
   int b0 = 0;
   for (auto &x : table)
       b0 += x.numbers * x.code.size();
   double res = (double)b / b0;
    return res;
int main()
   std::string s;
   std::ifstream f0("T:\\2kurs2sem\\InformTheor\\lab3\\TI_3.txt");
   f0 >> s;
   f0.close();
   int length = 2;
    std::vector<int> r = getSequencesOfNCharactersEach(s, length);
   clock_t start_time = clock();
    std::vector<character> res = theHuffmanMethod(r);
    clock_t end_time = clock();
    clock_t work_time = end_time - start_time;
    std::cout << "Time: " << (double)work_time;</pre>
    std::cout << "\nCompression ratio: " << getCompressionRatio(s, res);</pre>
    return 0;
```

Результат работы программы:

1) Метод Хаффмана

По 2 символа:

Time: 0

Compression ratio: 1

По 4 символа:

Time: 0

Compression ratio: 1.2458

По 8 символов:

Time: 0

Compression ratio: 2.165 2) Метод Шеннона-Фано

По 2 символа:

Time: 0

Compression ratio: 0.815842

По 4 символа:

Time: 0

Compression ratio: 1.17546

По 8 символов:

Time: 0

Compression ratio: 2.00097

# Время работы программы:

Метод построения кода	Количество символов в последовательности, взятой в качестве кодируемого символа		
	2	4	8
Метод Хаффмана	0	0	0
Метод Шеннона-Фано	0	0	0

# Коэффициент сжатия:

Метод построения кода	Количество символов в последовательности, взятой в качестве кодируемого символа		
	2	4	8
Метод Хаффмана	1	1,245	2.165
Метод Шеннона-Фано	0.815	1,175	2

#### Вывод:

По результатам, полученным в ходе работы программы и приведенным в таблицах выше, можно сделать следующие выводы. При анализе времени выполнения явного преобладания в скорости нет, так как язык С++ достаточно быстрый, но, если судить по сложности кода алгоритмов, алгоритм Шеннона-Фано будет эффективнее. Также мы сравнили коэффициенты сжатия. По данному показателю метод Хаффмана более эффективен, коэффициенты сжатия методом Хаффмана - больше, чем при использовании метода Шеннона-Фано. Сложность программной реализации обоих алгоритмов примерно одинаковая - средняя. Вручную же алгоритм Шеннона-Фано выполняется в более компактном и, следовательно, более удобном виде. Сложности выполнения этих алгоритмов вручнуютакже примерно одинаковы. Таким образом, можно сделать вывод, что алгоритм Хаффмана более эффективен, чем метод Шеннона-Фано. А значит, лучше использоватьметод Хаффмана, т.к. по скорости и простоте выполнения алгоритмы очень схожи.

#### №6

# Код программы:

```
#include "6.h"

bool areVectorsEqual(std::vector<int> a, std::vector<int> b)
{
    if (a.size() != b.size())
        return false;
    for (int i = 0; i < a.size(); i++)
        if (a[i] != b[i])
            return false;
    return true;
}

int getPosOfTheVector(const std::vector<character> &table, const
std::vector<int> &a)
{
    for (int i = 0; i < table.size(); i++)
        if (areVectorsEqual(table[i].code, a))
            return i;
    return -1;
}</pre>
```

```
std::string decoding(std::string codingS, std::vector<character> table, int
length)
    std::string res;
    std::vector<int> a;
    for (auto &x : codingS)
        a.push_back(x - '0');
        int pos = getPosOfTheVector(table, a);
        if (pos >= 0)
            std::vector<int> b(length, 0);
            b = getBinaryNumberNotation(b, 0, table[pos].symbol[0]);
            reverseVector(b);
            for (auto &y : b)
                res.push_back(y + '0');
            a.clear();
    }
    return res;
int main()
    std::string s;
    std::ifstream f0("T:\\2kurs2sem\\InformTheor\\lab3\\TI_3.txt");
    f0 >> s;
    f0.close();
    int length = 2;
    std::vector<int> r = getSequencesOfNCharactersEach(s, length);
    std::vector<character> res1 = theShannonFanoMethod(r);
    std::string sCode1 = replaceCharactersWithTheirCodes(res1, r);
    std::string s1 = decoding(sCode1, res1, length);
    std::vector<character> res2 = theHuffmanMethod(r);
    std::string sCode2 = replaceCharactersWithTheirCodes(res2, r);
    std::string s2 = decoding(sCode2, res2, length);
    if (s1 == s2)
        std::cout << "YES!!!\n";</pre>
    else
        std::cout << "NO!!\n";</pre>
    return 0;
```

Результат работы программы:

# По 2 символа:

YES!!!

# YES!!!

# По 8 символов:

# YES!!!

# **№**7

Расшифровка кода через сайт:

Ветер свистел, визжал, кряхтел и гудел на разные лады. То жалобным тоненьким голоском, то грубым басовым раскатом распевал он свою боевую песенку. Фонари чуть заметно мигали сквозь огромные белые хлопья снега, обильно сыпавшиеся на тротуары, на улицу, на экипажи, лошадей и прохожих.

**Вывод:** в ходе работы изучены возможности применения методов энтропийного кодирования для обработки двоичных последовательностей. Получены навыки написания и отладки программы составления кода для каждого символа сообщения методом Хаффмана и методом Шеннона-Фано, кодирования и декодирования двоичной последовательности. Сравнены время работы программы и коэффициенты сжатия.