МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ РЕСПУБЛИКИ БЕЛАРУСЬ

БЕЛОРУССКИЙ НАЦИОНАЛЬНЫЙ ТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ

***Факультет информационных технологий и робототехники***

Кафедра программного обеспечения информационных систем

и технологий

**Отчет по лабораторной работе № 8**

по дисциплине:” Алгоритмы и структуры данных”

на тему: ***“*** ***Алгоритмы статистического сжатия информации”***

Вариант 1

|  |  |  |
| --- | --- | --- |
| Исполнитель: |  | студент группы 10701322  Бородкин Д.В. |
| Преподаватель: |  | Мисякова В.А. |

Минск 2024

# Лабораторная работа №8. Алгоритмы статистического сжатия информации

# Цель работы: ****изучить основные алгоритмы статистическогоо сжатия информации и приобрести практические навыки работы с ними.****

# Задание:

1. Создать текстовый файл и записать в него свою фамилию, имя и отчество.
2. В соответствии с алгоритмами Шеннона-Фано и Хафмана закодировать текст.

(Создать деревья Шеннона-Фано и Хафмана)

1. Оценить коэффициент сжатия текста для каждого алгоритма.
2. Закодировать с помощью арифметического алгоритма свою фамилию.

### Алгоритм Шеннона-Фано

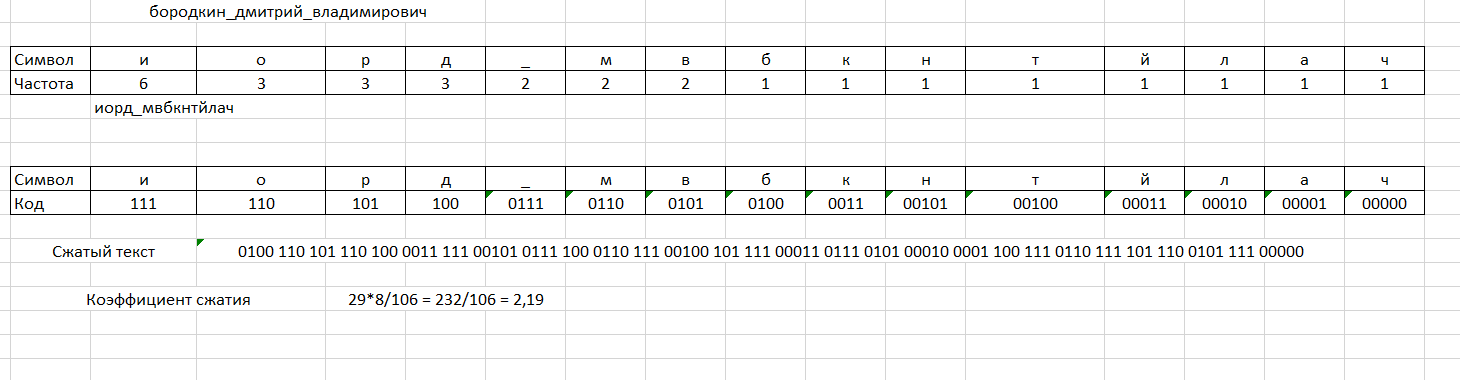


Рисунок 1 – Реализация алгоритма Шеннона-Фано в Excel

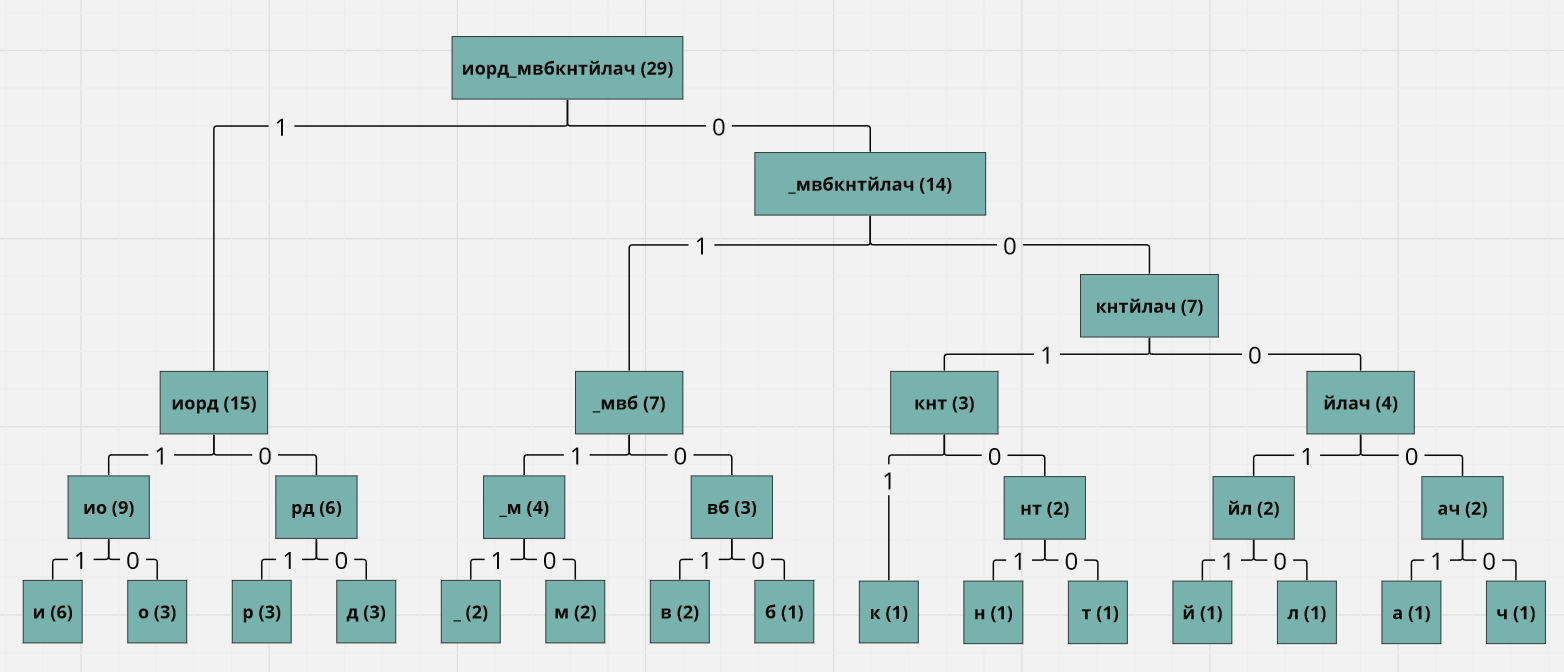


Рисунок 2 – Кодовое дерево

### Алгоритм Хаффмана

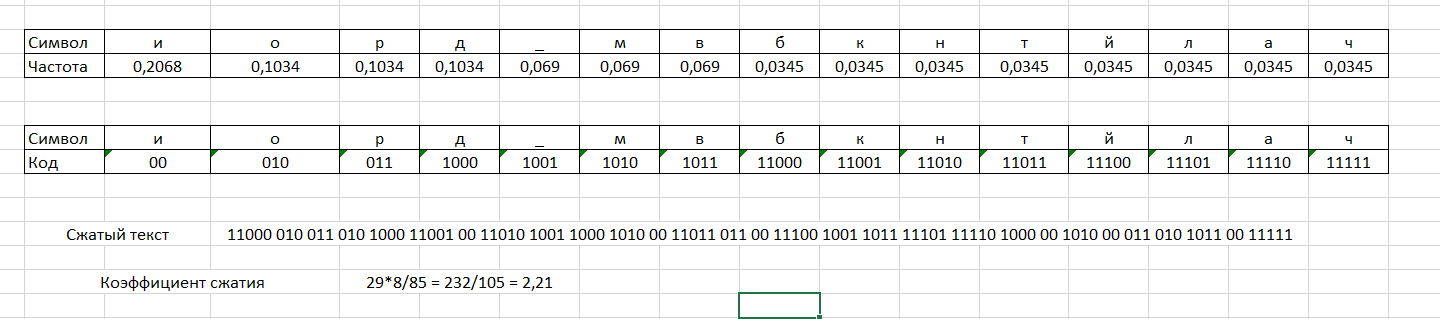


Рисунок 3 – Реализация алгоритма Хаффмана в Excel

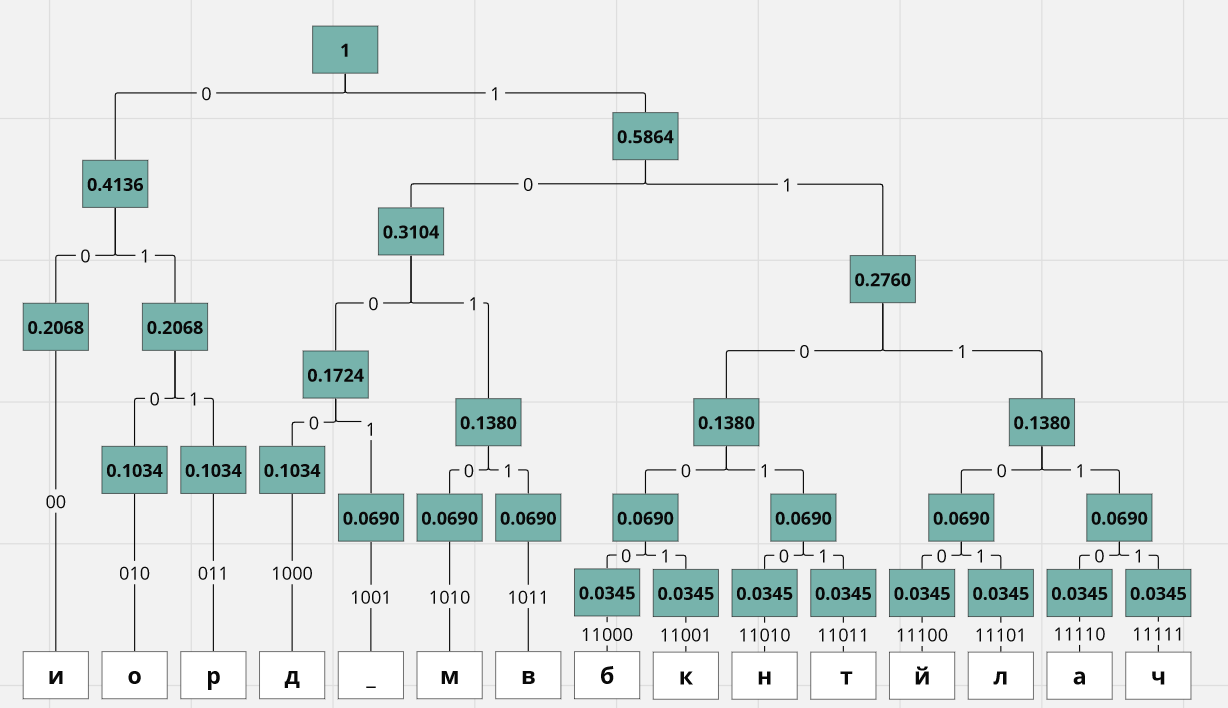


Рисунок 4 – Дерево кодирования

### Арифметическое кодирование

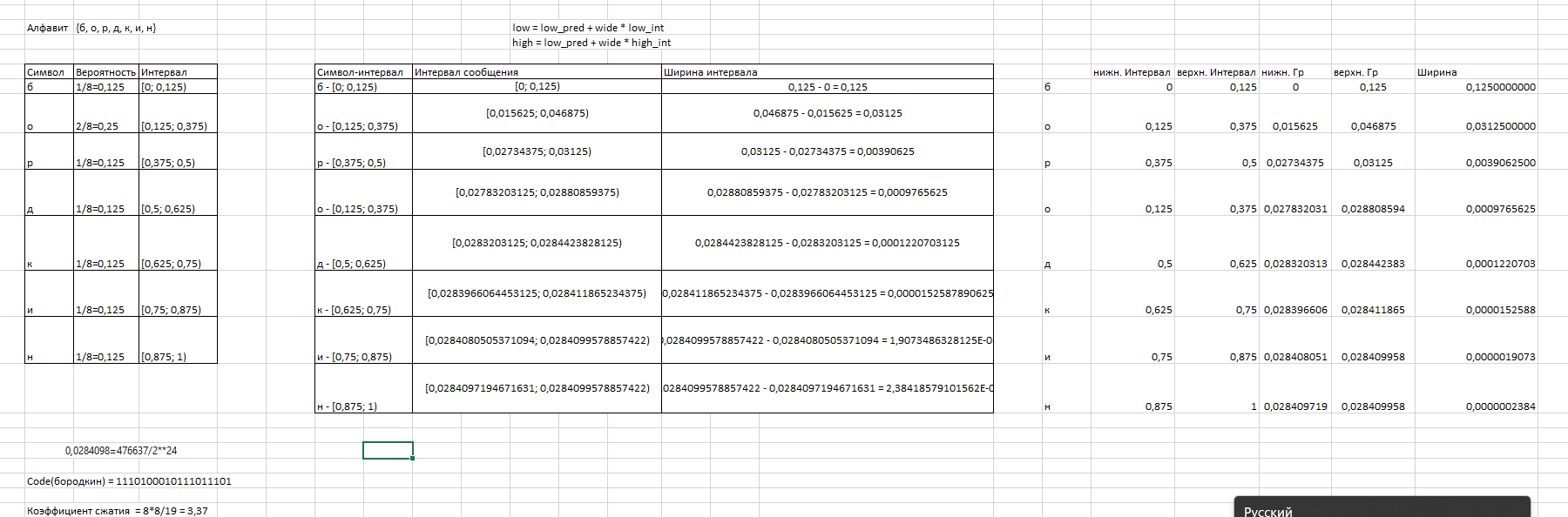


Рисунок 5 – Реализация арифметического кодирования в Excel

|  |  |  |
| --- | --- | --- |
| 0,0284098=476637/2\*\*24 | | |
|  |  |  |
| Code(бородкин) = 1110100010111011101 | | |
|  |  |  |
| Коэффициент сжатия = 8\*8/19 = 3,37 | | |

## Контрольные вопросы

1. Алгоритмы сжатия информации основаны на устранении избыточности, содержащейся в исходных данных. Простейшим примером избыточности является повторение в тексте фрагментов (например, слов естественного или машинного языка). Подобная избыточность обычно устраняется заменой повторяющейся последовательности ссылкой на уже закодированный фрагмент с указанием его длины. Другой вид избыточности связан с тем, что некоторые значения в сжимаемых данных встречаются чаще других. Сокращение объёма данных достигается за счёт замены часто встречающихся данных короткими кодовыми словами, а редких — длинными.
2. Коэффициент сжатия — основная характеристика алгоритма сжатия. Он определяется как отношение объёма исходных несжатых данных к объёму сжатых данных. То есть, если So — объём исходных данных, а Sc — объём сжатых данных, то коэффициент сжатия k определяется как k = So / Sc.
3. [В алгоритме **Шеннона-Фано**, сжатие производится путем замены повторяющихся последовательностей ссылкой на уже закодированный фрагмент с указанием его длины](https://ru.wikipedia.org/wiki/%D0%90%D0%BB%D0%B3%D0%BE%D1%80%D0%B8%D1%82%D0%BC_%D0%A8%D0%B5%D0%BD%D0%BD%D0%BE%D0%BD%D0%B0_%E2%80%94_%D0%A4%D0%B0%D0%BD%D0%BE). [Каждому символу присваивается уникальный код, и результатом работы алгоритма является закодированный текст](https://ru.stackoverflow.com/questions/598621/%D0%9B%D1%83%D1%87%D1%88%D0%B8%D0%B5-%D0%B0%D0%BB%D0%B3%D0%BE%D1%80%D0%B8%D1%82%D0%BC%D1%8B-%D1%81%D0%B6%D0%B0%D1%82%D0%B8%D1%8F-%D1%82%D0%B5%D0%BA%D1%81%D1%82%D0%B0). [Декодирование такого текста можно произвести, используя ту же самую структуру кодового дерева](https://ru.wikipedia.org/wiki/%D0%90%D0%BB%D0%B3%D0%BE%D1%80%D0%B8%D1%82%D0%BC_%D0%A8%D0%B5%D0%BD%D0%BD%D0%BE%D0%BD%D0%B0_%E2%80%94_%D0%A4%D0%B0%D0%BD%D0%BE).
4. В алгоритме **Хафмана**, сжатие достигается путем замены повторяющихся символов на более короткие кодовые слова. Каждому символу присваивается уникальный код, и результатом работы алгоритма является закодированный текст. Декодирование такого текста можно произвести, используя ту же самую структуру кодового дерева.
5. [В **арифметическом алгоритме**, сжатие информации производится путем представления каждого символа входного алфавита в виде некоторого интервала на числовой оси](https://mf.grsu.by/UchProc/livak/po/comprsite/theory_arithmetic.html). [Каждому символу присваивается уникальный код, и результатом работы алгоритма является закодированный текст](https://mf.grsu.by/UchProc/livak/po/comprsite/theory_arithmetic.html). [Декодирование такого текста можно произвести, используя ту же самую структуру кодового дерева](https://mf.grsu.by/UchProc/livak/po/comprsite/theory_arithmetic.html).

## Вывод: Сжатие данных является процессом, обеспечивающим уменьшение объема данных путем сокращения их избыточности.