

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)



ИНСТИТУТ ИНФОРМАЦИОННЫХ ТЕХНОЛОГИЙ И УПРАВЛЯЮЩИХ СИСТЕМ

Лабораторная работа №5
по дисциплине: Теория информации
тема: «Арифметическое кодирование»

Выполнил: ст. группы ПВ-223
Пахомов Владислав Андреевич

Проверили:
пр. Твердохлеб Виталий Викторович

Белгород 2024 г.

Лабораторная работа №5

Арифметическое кодирование

Цель работы: Исследовать особенности метода арифметического кодирования.

Задание 1. Построить обработчик, реализующий данный алгоритм арифметического кодирования. *Coder.java*

```
public static <T> ArithmeticEncodingTable<T> getArithmeticEncodingTable(List<T> input) {
    ArithmeticEncodingTable<T> aTable = new ArithmeticEncodingTable<>();
    List<TableElement<T>> segTable = getSegmentisedTable(input);

    for (TableElement<T> element : segTable) {
        aTable.addNewSymbol(element, (1.0 * element.amount) / input.size());
    }

    return aTable;
}

public static <T> AbstractMap.SimpleEntry<Double, Double> encodeArithmeticEncodingTable(List<T> input) {
    if (input.isEmpty()) return new AbstractMap.SimpleEntry<>(0., 0.);

    ArithmeticEncodingTable<T> table = getArithmeticEncodingTable(input);
    AbstractMap.SimpleEntry<Double, Double> answer =
        table.getProbabilitySpanBySymbol(input.get(0));

    for (int i = 1; i < input.size(); i++) {
        AbstractMap.SimpleEntry<Double, Double> currSymbol = table.getProbabilitySpanBySymbol(input.get(i));

        answer = new AbstractMap.SimpleEntry<>(
            answer.getKey() + (answer.getValue() - answer.getKey()) * currSymbol.getKey(),
            answer.getKey() + (answer.getValue() - answer.getKey()) * currSymbol.getValue());
    }

    return answer;
}
```

Main.java

```
package rchat.info.lab5;

import dn1.utils.text.table.TextTable;
import rchat.info.libs.Coder;

import javax.swing.event.TableModelListener;
import javax.swing.table.TableModel;
import java.io.BufferedReader;
import java.io.IOException;
import java.io.InputStreamReader;
import java.nio.file.Files;
import java.nio.file.Path;
import java.util.AbstractMap;
import java.util.ArrayList;
import java.util.Arrays;
```

```

import java.util.List;
import java.util.stream.Collectors;

public class Main {
    public static void main(String[] args) throws IOException {
        System.out.println("Введите сообщение: ");
        String finput;
        BufferedReader r = new BufferedReader(new InputStreamReader(System.in));
        finput = r.readLine();

        AbstractMap.SimpleEntry<Double, Double> ans =
        ↪  Coder.encodeArithmeticEncodingTable(Arrays.asList(finput.split(""))));

        double left = 0.0;
        double right = 1.0;
        double middle;
        int it = 0;

        while (true) {
            it++;
            middle = (right + left) / 2;

            if (middle > ans.getValue()) {
                right = middle;
            } else if (middle < ans.getKey()) {
                left = middle;
            } else break;
        }

        System.out.println("Интервал: " + ans);
        System.out.println("Число, характеризующее интервал: " + middle);
        System.out.println("Количество бит: " + it);
        System.out.println("Код: ");
        while (middle != 0) {
            System.out.print(((int)(middle * 2) == 1 ? "1" : "0"));
            middle *= 2;
            if (middle >= 1)
                middle -= 1;
        }
    }
}

```

Задание 2. Построить код для сообщения, содержащего строку панграммы «в чащах юга жил бы цитрус? Да, но фальшивый экземпляр!».

Введите сообщение:

в чащах юга жил бы цитрус? Да, но фальшивый экземпляр!

Интервал: 0.4505322663570715=0.4505322663570715

Число, характеризующее интервал: 0.4505322663570715

Количество бит: 54

Код:

01110011010101100001010100100101110010111101111110001

Количество данных по сравнению с кодированием по методу Шеннона Фано уменьшилось, $54 < 258$.

Коэффициент сжатия: 8

Задание 3. Построить код для сообщения, содержащего строку «Victoria nulla est, Quam quae confessos animo quoque subjugat hostes»

Введите сообщение:

Victoria nulla est, Quam quae confessos animo quoque subjugat hostes

Интервал: $0.9966629048585032=0.9966629048585032$

Число, характеризующее интервал: 0.9966629048585032

Количество бит: 52

Код:

1111111100100101010011001101010110000000111011011111

Количество данных по сравнению с кодированием по методу Шеннона Фано уменьшилось, $52 < 278$.

Коэффициент сжатия: $10\frac{6}{13}$

Вывод: в ходе лабораторной работы исследовали особенности метода арифметического кодирования. Эффективность сжатия алгоритма сильно превосходит алгоритм Шеннона Фано.