

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 № 2
по дисциплине: Теория информации
тема: «Исследование кодов Шеннона-Фано»

Выполнил: ст. группы ПВ-223

Игнатьев Артур Олегович

Проверил:

Твердохлеб Виталий Викторович

Белгород 2024г.

Лабораторная работа №2

«Исследование кодов Шеннона-Фано»

Цель работы: исследовать кодирование по методу Шеннона-Фано.
Научиться оценивать эффективности кода..

Решение задач:

1. Построить код для сообщения, содержащего строку панграммы «в чашах юга жил бы цитрус? Да, но фальшивый экземпляр!». Для полученного кода рассчитать показатели коэффициента сжатия и дисперсии.

Символ	Вероятность	Этапы						Код
		I	II	III	IV	V	VI	
пробел	0,166666667	1	1	1				111
А	0,092592593			0				110
Л	0,055555556		0	1	1			1011
И	0,055555556				0			1010
В	0,037037037			0	1	1		10011
Ы	0,037037037					0		10010
Р	0,037037037				0	1		10001
П	0,018518519					0		10000
?	0,018518519	1	1	1	1	1		01111
О	0,018518519					0	1	011101
Н	0,018518519						0	011100
М	0,018518519				0	1		01101
К	0,018518519					0	1	011001
И	0,018518519					0	0	011000
З	0,018518519			0	1	1		01011
Ж	0,018518519					0	1	010101
Е	0,018518519						0	010100
Г	0,018518519				0	1	1	010011
Б	0,018518519					0	0	010010
,	0,018518519					0	1	010001
!	0,018518519					0	0	010000
Д	0,018518519		0	1	1	1		00111
Я	0,018518519					0	1	001101
Ю	0,018518519						0	001100
Э	0,018518519				0	1	1	001011
Ь	0,018518519						0	00101
Щ	0,018518519					0	1	001001
Ш	0,018518519					0	0	001000
Ч	0,018518519			0	1	1		00011
Ц	0,018518519					0	1	000101
Х	0,018518519						0	000100
Ф	0,018518519				0	1	1	000011
У	0,018518519						0	000010
Т	0,018518519					0	1	000001
С	0,018518519					0	0	000000

Код сообщения: 10011 111 00011 110 001001 110 000100 111 001100 010011
110 111 010101 1010 1011 111 010010 10010 111 000101 1010 000001 10001
000010 000000 01111 111 00111 110 010001 111 011100 011101 111 000011 110
1011 001010 001000 1010 10011 10010 011000 111 001011 011001 01011 010100
01101 10000 1011 001101 10001 010000

$$B = 54 \cdot 8 = 432$$

$$B' = 258$$

$$K_{comp} = \frac{B}{B'} = \frac{432}{258} = 1\frac{29}{43}$$

$$l = 4.7778$$

$$\delta = 1.5062$$

2. Построить код для сообщения, содержащего строку «Victoria nulla est, Quam quae confessos animo quoque subjugat hostes» Для полученного кода рассчитать показатели коэффициента сжатия и дисперсии.

Символ	Вероятность	Этапы							Код	
		I	II	III	IV	V	VI	VII		
пробел	9	1	1	1					111	
u	7			0					110	
s	7			1					101	
o	6		0	0	1				1001	
a	6			0	0				1000	
e	5									
t	4	0	1	1	1				0111	
q	3			0	0				0110	
n	3			0	1				0101	
i	3			0				0100		
m	2		1	1	1	1			00111	
l	2				0	0			00110	
c	2	0			1			00101		
V	1	0	0	1	0			00100		
r	1				1	1			000111	
Q	1				0	0			000110	
,	1				0	1			000101	
j	1		0	0	1	0			000100	
h	1					1	1			000011
g	1					0	0			000010
f	1				1	1			000001	
b	1				0	0	1	1		0000001
							0	0	1	

Код сообщения: 000111 00111 00100 0110 1001 000110 00111 1000 111 0100
 110 00101 00101 1000 111 0111 101 0110 000100 111 000101 110 1000 00110
 111 0101 110 1000 0111 111 00100 1001 0100 0000001 0111 101 101 1001 101
 111 1000 0100 00111 00110 1001 111 0101 110 1001 0101 110 0111 111 101 110
 0000000 000011 110 000001 1000 0110 111 000010 1001 101 0110 0111 101

$$B = 68 \cdot 8 = 544$$

$$B' = 278$$

$$K_{comp} = \frac{B}{B'} = \frac{544}{278} = 1 \frac{133}{139}$$

$$l = 4.0882$$

$$\delta = 0.5515$$

3. Построить консольное приложение, реализующее процесс кодирования по методу Шеннона-Фано (с возможностью расчета коэффициента сжатия и дисперсии).

```
package labs.lab2;

import java.io.BufferedReader;
import java.io.IOException;
import java.io.InputStreamReader;
import java.util.*;
import java.util.stream.Collectors;

public class Main {
    public static class TableElement {
        char symbol;
        int amount;
        List<Boolean> code;
        public TableElement(char symbol) {
            this.symbol = symbol;
            this.amount = 1;
            this.code = new ArrayList<>();
        }
    }

    public static List<TableElement> schennonFano(String input) {
        List<TableElement> table = new ArrayList<>();
        for (char symbol : input.toCharArray()) {
            Optional<TableElement> result = table.stream().filter((el) ->
el.symbol == symbol).findAny();
            if (result.isPresent()) {
                result.get().amount++;
            } else {
                table.add(new TableElement(symbol));
            }
        }
        table.sort(Comparator.comparingInt(o -> o.amount));
        Collections.reverse(table);
        schennonFano(table, 0, table.size());
        return table;
    }

    public static void schennonFano(List<TableElement> table, int beginIndex,
int endIndex) {
        if (endIndex - beginIndex <= 1) return;
        if (endIndex - beginIndex == 2) {
            table.get(beginIndex).code.add(true);
            table.get(beginIndex + 1).code.add(false);
            return;
        }
        int separateIndex = getSeparateIndex(table, beginIndex, endIndex);
        for (int i = beginIndex; i < endIndex; i++) {
            if (i < beginIndex + separateIndex) {
                table.get(i).code.add(true);
            } else {
                table.get(i).code.add(false);
            }
        }
        schennonFano(table, beginIndex, separateIndex);
        schennonFano(table, separateIndex, endIndex);
    }

    private static int getSeparateIndex(List<TableElement> table, int
beginIndex, int endIndex) {
        int sum = 0;
        for (int i = beginIndex; i < endIndex; i++) {
```

```

        sum += table.get(i).amount;
    }
    int sumBefore = table.get(beginIndex).amount;
    int sumAfter = sum - sumBefore;
    int separateIndex = beginIndex + 1;
    while (separateIndex < endIndex - 1 && sumAfter - table.get(sepa-
rateIndex).amount -
        (sumBefore + table.get(separateIndex).amount) > 0 ) {
        sumAfter -= table.get(separateIndex).amount;
        sumBefore += table.get(separateIndex).amount;
        separateIndex++;
    }
    if (Math.abs(sumBefore - sumAfter) > Math.abs(sumAfter - ta-
ble.get(separateIndex).amount -
        (sumBefore + table.get(separateIndex).amount))) {
        separateIndex++;
    }
    return separateIndex;
}
}
public static void main(String[] args) throws IOException {
    System.out.println("Введите сообщение: ");
    String input;
    BufferedReader r = new BufferedReader(new InputStreamReader(Sys-
tem.in));
    input = r.readLine();
    System.out.println();
    List<TableElement> table = schennoFano(input);
    System.out.println("Таблица: ");
    for (TableElement element : table) {
        System.out.print(element.symbol);
        System.out.print(" " + element.amount + " ");
        for (int i = 0; i < element.code.size(); i++) {
            System.out.print(element.code.get(i) ? "1" : "0");
        }
        System.out.println();
    }
    System.out.println();
    System.out.println("Закодированное сообщение: ");
    String code = input;
    int sum = 0;
    for (TableElement element : table) {
        String elementCode = element.code.stream().map((el) -> el ? "1" :
"0").collect(Collectors.joining(""));
        code = code.replace(" " + element.symbol, elementCode);
        sum += element.amount;
    }
    System.out.println(code);
    System.out.println();
    int codedLength = code.length();
    int uncodedLength = input.length() * 8;
    System.out.println("Коэффициент сжатия: " + 1.0 * uncodedLength /
codedLength);
    System.out.println();
    double midLen = 0;
    for (TableElement element : table) {
        String elementCode = element.code.stream().map((el) -> el ? "1" :
"0").collect(Collectors.joining(""));
        midLen += elementCode.length() * (1.0 * element.amount / sum);
    }
    System.out.println("Средняя длина: " + midLen);
    System.out.println();
    double delta = 0;
    for (TableElement element : table) {
        String elementCode = element.code.stream().map((el) -> el ? "1" :

```

```
"0").collect(Collectors.joining(""));
        delta += (1.0 * element.amount / sum) * (elementCode.length() -
midLen) * (elementCode.length() - midLen);
    }
    System.out.println("Дисперсия: " + delta);
}
}
```

4. Получить кодовые представления сообщений из пунктов 1 и 2 задания по методу Хаффмана. Сравнить полученные результаты с методом Шеннона-Фано по показателям сжатия и дисперсии. Сделать соответствующие выводы.

Результаты выполнения программы для задания 1:

Введите сообщение:

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

Таблица:

9	111
а	5 110
л	3 1011
и	3 1010
р	2 10111
ы	2 10110
в	2 10011
!	1 10010
я	1 01111
п	1 011111
м	1 011110
е	1 01111
э	1 011111
к	1 011110
э	1 01111
й	1 011111
ш	1 011110
ь	1 011111
ф	1 011110
о	1 011111
н	1 011110
,	1 01111
д	1 011111
?	1 011110
с	1 011111
у	1 011110
т	1 011111
ц	1 011110
б	1 01111
ж	1 001111
г	1 001110
ю	1 001111
х	1 001110
щ	1 001111
ч	1 001110

Закодированное сообщение:

10011111001110110001111110001110111001111001110110111001111101010111110111110
11011101111010100111111011101111001111101111011101111111001111111011110011111
111011110110101101111110111101010100111011001111111101111011110011111011110111
1001111110110111111011110010

Коэффициент сжатия: 1.6744186046511629

Средняя длина: 4.777777777777776

Дисперсия: 1.5061728395061729

Результаты выполнения программы для задания 2:

```
Введите сообщение:
Victoria nulla est, Quam quae confessos animo quoque subjugat hostes

Таблица:
 9 111
s 7 110
u 7 101
a 6 1011
o 6 1010
e 5 0111
t 4 0110
q 3 0111
n 3 0110
i 3 01111
m 2 01110
l 2 01111
c 2 01110
h 1 011111
g 1 001110
j 1 001111
b 1 001110
f 1 001111
Q 1 001110
, 1 001111
r 1 0011111
v 1 0011110

Закодированное сообщение:
00111100111101110011010100011111011111011111011010101111011111011111011111001
10001111111001110101101101110111011110110110111111011101010011000111101111101
1010101101111011011001111011101010111011110110100111101011111110101001110001
11110100111010110110111011111101011001100111110

Коэффициент сжатия: 1.9568345323741008

Средняя длина: 4.088235294117648

Дисперсия: 1.1392733564013837
```

Вывод: в ходе лабораторной работы исследовали кодирование по методу Шеннона-Фано. Научиться оценивать эффективности кода. Выполнили реализацию метода кодирования по методу Шеннона-Фано в виде консольного приложения.