

ГУАП

КАФЕДРА №

ОТЧЕТ
ЗАЩИЩЕН С ОЦЕНКОЙ _____

ПРЕПОДАВАТЕЛЬ

Гений науки, к.г.н.				Суетина Т. А.
должность, уч. степень, звание		подпись, дата		инициалы, фамилия

ОТЧЕТ О ЛАБОРАТОРНОЙ РАБОТЕ №1

Вариант 5

по курсу: Энтропийные алгоритмы сжатия информации

РАБОТУ ВЫПОЛНИЛ

СТУДЕНТ ГР. №	4128			Анонимный Н. Н.
			подпись, дата	инициалы, фамилия

Санкт-Петербург 2024

СОДЕРЖАНИЕ

1	Введение	3
1.1	Цель лабораторной работы	3
1.2	Задание	3
2	Выполнение работы	4
2.1	Метод Хаффмана	4
2.2	Метод Шенона-Фано	5
2.3	Арифметическое кодирование	5
2.4	Алгоритм LZW	7
3	Вывод	8

1 Введение

1.1 Цель лабораторной работы

1.2 Задание

Выполнить сжатие текста 4 способами:

- Метод Хаффмана;
- Метод Шенона-Фано;
- Арифметическим кодированием;
- Алгоритмом LZW.

Для каждого метода рассчитать коэффициент сжатия текста.

Вариант 5: ШОРОХ ОТ ДУБКА КАК БУДТО ХОРОШ

2 Выполнение работы

Для начала проанализируем текст.

Таблица 2.1 - фигура

Буква	Ш	О	Р	Х	space	Д	У	Б	К	А	Т
Кол-во	2	6	2	2	5	2	2	2	3	2	2

Всего букв: 30

2.1 Метод Хаффмана

Таблица 2.2 - Решение методом Хаффмана

Буква	О	space	К	Ш	Р	Х	Д	У	Б	А	Т
Частота	6	5	3	2	2	2	2	2	2	2	2
	1			0							
	1	0		1				0			
		1	0	1		0		1		0	
				1	0	1	0	1	0	1	0
ИТОГ	11	101	100	0111	0110	0101	0100	0011	0010	0001	0000

Итоговый текст:

[0111 11 0110 11 0101]101[11 0000]101[0100 0011 0010 100 0001]101
[100 0001 100]101[0010 0011 0100 0000 11]101[0101 11 0110 11 0111]

Коэффициент кодирования: $100/120 = 0.83$

2.2 Метод Шенона-Фано

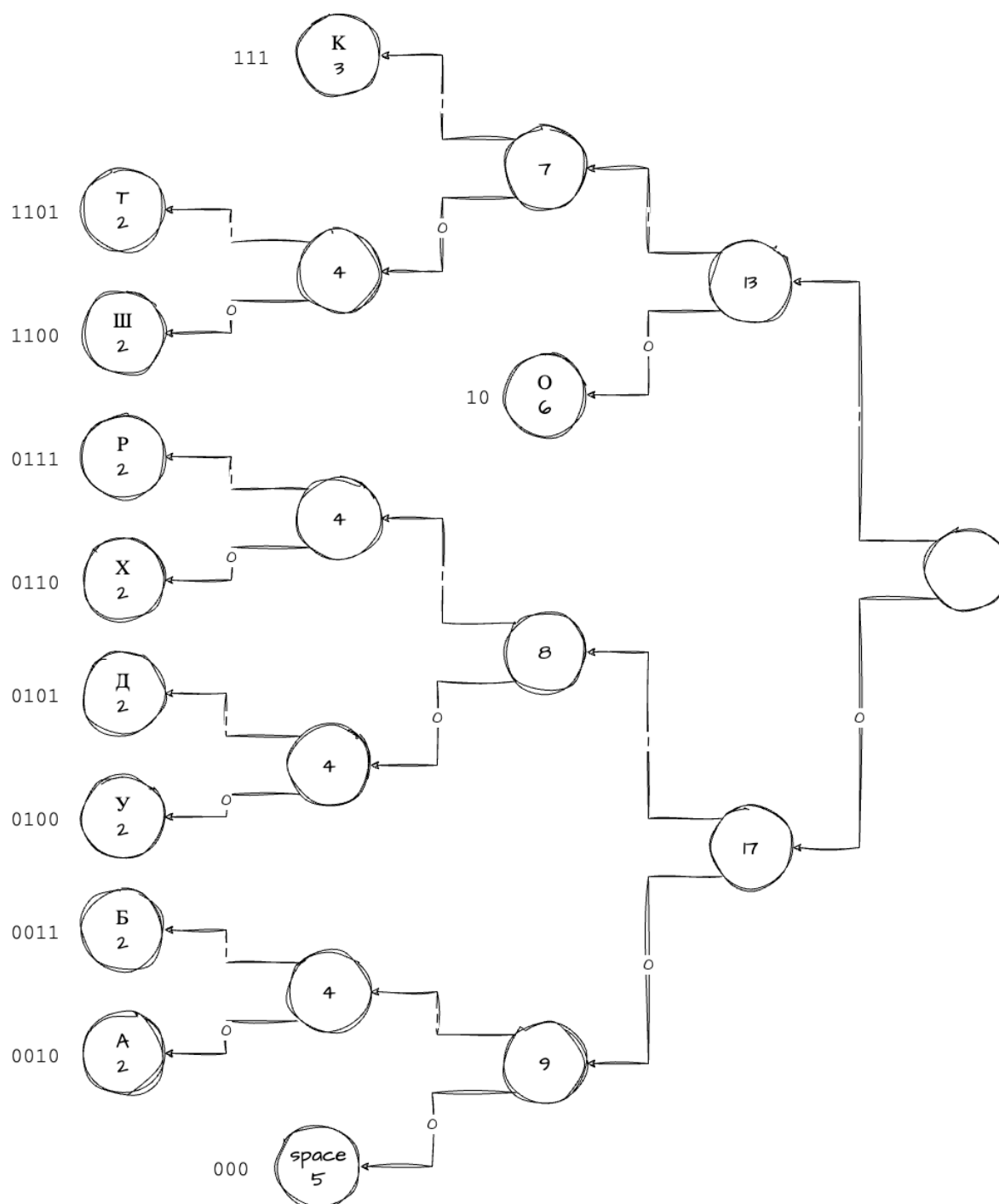


Рисунок 2.1 - Граф для метода Шенона-Фано

Итоговый текст:

[1100 10 0111 10 0110]000[10 1101]000[0101 0100 0011 111 0010]000
[111 0010 111]000[0011 0100 0101 1101 10]000[0110 10 0111 10 1100]

Коэффициент кодирования: $100/120 = 0.83$

2.3 Арифметическое кодирование

Бв	Ш	О	Р	Х	space	Д	У	Б	К	А
Ит.	$\frac{2}{28}$	$\frac{8}{28}$	$\frac{10}{28}$	$\frac{12}{28}$	$\frac{17}{28}$	$\frac{19}{28}$	$\frac{21}{28}$	$\frac{23}{28}$	$\frac{26}{28}$	$\frac{28}{28}$
Н.	$\frac{0}{28}$	$\frac{2}{28}$	$\frac{8}{28}$	$\frac{10}{28}$	$\frac{12}{28}$	$\frac{17}{28}$	$\frac{19}{28}$	$\frac{21}{28}$	$\frac{23}{28}$	$\frac{26}{28}$

```

Character: Ш, Interval: (0, 1/15)
Character: О, Interval: (1/225, 4/225)
Character: Р, Interval: (1/125, 2/225)
Character: О, Interval: (136/16875, 139/16875)
Character: Х, Interval: (137/16875, 686/84375)
Character: , Interval: (3427/421875, 20567/2531250)
Character: О, Interval: (61687/7593750, 6169/759375)
Character: Т, Interval: (68543/8437500, 616889/75937500)
Character: , Interval: (3084439/379687500, 4626661/569531250)
Character: Д, Interval: (55519921/6834375000, 18506641/2278125000)
Character: У, Interval: (23133301/2847656250, 416399419/51257812500)
Character: Б, Interval: (12491982563/1537734375000, 2498396513/307546875000)
Character: К, Interval: (18737973847/2306601562500, 187379738473/23066015625000)
Character: А, Interval: (234224673091/28832519531250, 187379738473/23066015625000)
Character: , Interval: (2342246730911/288325195312500, 28106960770937/3459902343750000)
Character: К, Interval: (168641764625617/20759414062500000, 8432088231281/1037970703125000)
Character: А, Interval: (843208823128099/103797070312500000, 8432088231281/1037970703125000)
Character: К, Interval: (5059252938768599/622782421875000000, 12648132346921499/1556956054687500000)
Character: , Interval: (126481323469214981/15569560546875000000, 28106960770936663/3459902343750000000)
Character: Б, Interval: (303555176326115959/3736694531250000000, 1517775881630579797/186834726562500000000)
Character: У, Interval: (3794439704076449491/467086816406250000000, 5691659556114674237/700630224609375000000)
Character: Д, Interval: (341499573366880454209/42037813476562500000000, 113833191122293484737/1401260449218750000000)
Character: Т, Interval: (160077925015725212911/19705225067138671875000, 2561246800251603406577/31528360107421875000000)
Character: О, Interval: (38418702003774051098641/4729254016113281250000000, 9604675500943512774661/1182313504028320312500000)
Character: , Interval: (192093510018870255493211/23646270080566406250000000, 128062340012580170328809/1576418005371093750000000)
Character: Х, Interval: (1152561060113221532959271/141877620483398437500000000, 16007792501572521291101/1970522506713867187500000)
Character: О, Interval: (8644207950849161497194533/1064082153625488281250000000, 17288415901698322994389069/212816430725097656250000000)
Character: Р, Interval: (14407013251415269161990889/1773470256042480468750000000, 17288415901698322994389067/212816430725097656250000000)
Character: О, Interval: (1296631192627374224579180011/159612323043823242187500000000, 648315596313687112289590007/7980616152191162109375000000)
Character: Ш, Interval: (1296631192627374224579180011/159612323043823242187500000000, 270131498464036296787329169/3325256730079650878906250000)
Final interval: 0.008123628350872197329761728237972960156458303867848849348266122787823419923959139374770115972344954096, 0.008123628350872197329761728239225996239048793472248099014310119371214888029436265680941940178193426534
Interval cut : 0.0081236283508721973297617282379, 0.0081236283508721973297617282392
Prefix : 0.00812362835087219732976172823

```

Рисунок 2.2 - Результат арифметического кодирования

Видно, что получившийся полуинтервал имеет начало 0.0.007491234978280553441395982224376491768148032361088748495779293270670468164890615017316007369538876504 и конец 0.007491234978280553441395982224391606091525613885464659445273216987363890074114548014936179948312863938.

Исходя из рисунка 2.2, можно сделать вывод, что сообщение можно зашифровать числом $0.00749123497828055344139598222438_{10} = 0.000000001111010101111_2$, то есть 000000001111010101111_2 .

Коэффициент кодирования: $20/120 = 0.16$

2.4 Алгоритм LZW

```
Ш: 0   Д: 1   : 2   Р: 3
О: 4   Б: 5   А: 6   К: 7
Х: 8   Т: 9   У: 10  ШО: 11
ОР: 12  РО: 13  ОХ: 14  Х : 15
О: 16  ОТ: 17  Т : 18  Д: 19
ДУ: 20  УБ: 21  БК: 22  КА: 23
А : 24  К: 25  КАК: 26  К : 27
Б: 28  БУ: 29  УД: 30  ДТ: 31
ТО: 32  О : 33  Х: 34  ХО: 35
ОРО: 36  ОШ: 37  Encoded data: [0, 4, 3, 4, 8, 2, 4, 9, 2, 1, 10, 5, 7,
6, 2, 23, 7, 2, 5, 10, 1, 9, 4, 2, 8, 12, 4, 0]
Size of the encoded data in bits: 140
```

Рисунок 2.3 - Результат работы LZW

Коэффициент кодирования: $140/120 = 1.16$

3 Вывод