



Міністерство освіти і науки України  
Національний технічний університет України  
«Київський політехнічний інститут імені Ігоря Сікорського»  
Фізико-технічний інститут

КРИПТОГРАФІЯ  
КОМП'ЮТЕРНИЙ ПРАКТИКУМ №1  
**Експериментальна оцінка ентропії на символ джерела відкритого тексту**

**Виконали:**

Студенти групи ФБ-73  
Мухамедзянов Артем  
Лень Олександр

Київ-2019

## Мета роботи

Засвоєння понять ентропії на символ джерела та його надлишковості, вивчення та порівняння різних моделей джерела відкритого тексту для наближеного визначення ентропії, набуття практичних навичок щодо оцінки ентропії на символ джерела.

## Порядок виконання роботи

0. Уважно прочитати методичні вказівки до виконання комп'ютерного практикуму.

1. Написати програми для підрахунку частот букв і частот біграм в тексті, а також підрахунку  $H_1$  та  $H_2$  за безпосереднім означенням. Підрахувати частоти букв та біграм, а також значення  $H_1$  та  $H_2$  на довільно обраному тексті російською мовою достатньої довжини (щонайменше 1Мб), де імовірності замінити відповідними частотами. Також одержати значення  $H_1$  та  $H_2$  на тому ж тексті, в якому вилучено всі пробіли.

2. За допомогою програми CoolPinkProgram оцінити значення  $(10) H$ ,  $(20) H$ ,  $(30) H$ .

3. Використовуючи отримані значення ентропії, оцінити надлишковість російської мови в різних моделях джерела.

## Хід роботи, опис труднощів, що виникали, та шляхи їх розв'язання

Першим кроком до виконання комп'ютерного практикуму було отримання важливої інформації у методичці та на лекціях, щодо того, що таке ентропія, як її рахувати та для чого вона використовується. Отримавши завдання практикуму, перш за все, ми вирішили, на якій мові будемо писати нашу програму та які основні засоби ми використаємо. Розбивши наше завдання по етапам, ми прийняли їх втілювати у життя, а саме написання циклів підрахунку частот букв. Для виконання нашої роботи ми використовували знайому всім і діючу методику вирішення задач будь-якої складності «Step by step». Але, як, мабуть, і у всіх, у нас почали виникати труднощі в написанні робочого коду, бо кількість помилок на перших етапах вражала ⇒

Найбільшою проблемою стало написання циклу запису підрахунку «не перекрестних» біграм до txt файлу. Вирішенням цієї проблеми стало «Гуглення» розв'язку близько двох годин, бо всі запропоновані засоби вирішення подібної проблеми, як у нас, або не працювали, або працювали некоректно. Але, з часом, ми знайшли потрібний спосіб вирішення і здивувалися, як самі до цього не додумались.

## ***Код программы***

```
#include "locale.h"
#include "stdlib.h"
#include <fstream>
#include <ctype.h>
#include <algorithm>
    #include <iostream>
#include<math.h>
#include<string.h>
using namespace std;

    char AlphaZ[]="абвгдеёжзийклмнопрстуфхцчщщъыьэюя ";
int sizeAlpha = 34;
int mas, cc=0;
int **bigram;
long double count_monograms, count_bigrams = 0;
long double entropy = 0.0;
long double chastota = 0.0;
int monograms[34]={0};
void Sort()
{
    int temp;
    char ctemp;
    for (int i = 0; i < sizeAlpha; i++)
    {
        for (int j = 0; j < sizeAlpha - i - 1; j++) {
            if (monograms[j] < monograms[j + 1]) {

                temp = monograms[j];
                monograms[j] = monograms[j + 1];
                monograms[j + 1] = temp;

                ctemp = AlphaZ[j];
                AlphaZ[j] = AlphaZ[j + 1];
                AlphaZ[j + 1] = ctemp;

            }
        }
    }
}
```

```
}
```

```
}
```

```
void CountMonograms(char *text, char *AlphaZ)
```

```
{
```

```
    char temp;
```

```
    while(*text)
```

```
    {
```

```
        temp = tolower(*text);
```

```
        for(int i=0;i<sizeAlpha;i++)
```

```
        {
```

```
            if(temp==AlphaZ[i])
```

```
            {
```

```
                monograms[i]++;
```

```
                count_monograms++;
```

```
            }
```

```
        }
```

```
        text++;
```

```
    }
```

```
}
```

```
void CountBigrams(char *text, char *AlphaZ)
```

```
{
```

```
    char prev=0;
```

```
    while(*text)
```

```
    {
```

```
        char curr = tolower(*text);
```

```
        char *prev_ = find(AlphaZ, AlphaZ+mas, prev);
```

```
        char *curr_ = find(AlphaZ, AlphaZ+mas, curr);
```

```
        if( cc%2==0    && curr_ != AlphaZ+mas && prev_ != AlphaZ+mas)
```

```
        {
```

```
            bigram[prev_ - AlphaZ][curr_ - AlphaZ]++;
```

```
            count_bigrams++;
```

```
        }
```

```
        cc++;
```

```

        prev = curr;
        text++;
    }
}

void SaveBigrams(char *filename, char *AlphaZ)
{
    ofstream ofs;
    ofs.open (filename, std::ofstream::out);
    ofs << "      ";
    for( int i = 0; i < mas; i++ )
        ofs << AlphaZ[i] << "      ";
    ofs << endl;

    for( int j = 0; j < mas; j++ )
    {
        ofs << AlphaZ[j] << " ";
        for( int i = 0; i < mas; i++ )
        {
            ofs.width(4);
            ofs << bigram[i][j] << " ";
        }
        ofs << endl;
    }
    ofs.close();
}

long double EntropyMonograms()
{
    chastota, entropy = 0;
    Sort();

    for(int i=0;i<sizeAlpha;i++)
    {
        chastota = (monograms[i]/count_monograms);

        if(chastota!=0)

```

```

        {
            cout<<AlphaZ[i]<<" "<<monograms[i]<<" "<<chastota<<endl;
            entropy -=  chastota  * (log(chastota)/log(2));
        }
    }
return entropy;
}

```

```

long double EntropyBigrams()
{
    chastota, entropy = 0;
    for(int i = 0; i < mas; i++)
    {
        for(int j = 0; j<mas; j++)
        {
            chastota =  (bigram[i][j]/count_bigrams);
            if(chastota!=0)
            {
                entropy -=  chastota  * (log(chastota)/log(2));
            }
        }
    }
return  entropy;
}

```

```

int main(int argc, char* argv[])
{
    setlocale(LC_ALL,"rus");
    string str;
    ifstream file;
    file.open ( "text.txt" );
    getline ( file, str, '\0' );
    file.close();
    char *input_text = new char[str.length()+1];
    strcpy(input_text, str.c_str());

    mas = sizeof(AlphaZ) - 1;

```

```

bigram = new int* [mas];

for(int i = 0; i < mas; i++)
{
    bigram[i] = new int[mas];
    for(int j = 0; j<mas; j++)
        bigram[i][j] = 0;
}

CountBigrams( input_text, AlphaZ );

SaveBigrams( "bigrams неперекрёстные.txt", AlphaZ);

CountMonograms( input_text, AlphaZ );

cout<<endl;

cout<<"H1 : "<< EntrophyMonograms() <<endl;

cout<<"H2 : "<< EntrophyBigrams() <<endl;

cout<<fixed<<"Всего символов в тексте : "<<int(count_monograms)<<endl;

cout<<fixed<<"Всего биграмм в тексте : "<<int(count_bigrams)<<endl;

    return 0;
}

```

**Скріншоти до результатів розової програмки :**

H(10).jpg

Лабораторная работа №1

Произвольная часть текста:  
му\_что\_в\_

Использованные буквы:

Порядок n-граммы:  
5 символов  
10 символов  
15 символов  
20 символов  
25 символов  
30 символов  
35 символов  
40 символов  
45 символов  
50 символов

Введенный символ:

Символ по счету:

Номер эксперимента: 51

Неравенство для энтропии:  
 $2,79604428508995 < H < 3,52237881293525$

Двоичная таблица угаданных символов:

00000100000000000000000000000000
00000000000100000000000000000000
10000000000000000000000000000000
10000000000000000000000000000000
10000000000000000000000000000000

Вероятности:

q[1] = 0,28
q[2] = 0,14
q[3] = 0,08
q[4] = 0,02
q[5] = 0,1
q[6] = 0,06
q[7] = 0,02
q[8] = 0,02
q[9] = 0
q[10] = 0,06
q[11] = 0,02
q[12] = 0,04
q[13] = 0,04
q[14] = 0,02
q[15] = 0
q[16] = 0
q[17] = 0,02
q[18] = 0,02
q[19] = 0,02
q[20] = 0
q[21] = 0
q[22] = 0
q[23] = 0
q[24] = 0
q[25] = 0,02
q[26] = 0
q[27] = 0
q[28] = 0
q[29] = 0
q[30] = 0,02
q[31] = 0
q[32] = 0

Поле ввода символов:

Продолжить Другой

Строка состояния:

Лабораторная работа №1

Лабораторная работа №1

**Произвольная часть текста:**

oe\_ему\_to\_не\_успеете\_выи\_сло

---

**Использованные буквы:**

---

  

<b>Порядок n-граммы:</b> 5 символов 10 символов 15 символов 20 символов 25 символов <u>30 символов</u> 35 символов 40 символов 45 символов 50 символов	<b>Введенный символ:</b> _____  <b>Символ по счету:</b> _____  <b>Номер эксперимента:</b> 51	<b>Неравенство для энтропии:</b> <b>1.568498949333992 &lt; H &lt; 2.34093396247092</b>
--	--	---

  

<b>Поле ввода символов:</b> <input type="text"/> <div> <span>Продолжить</span> <span>Другой</span> </div>	<b>Двоичная таблица угаданных символов:</b> <pre> 1000000000000000000000000000000000 ^ 1000000000000000000000000000000000  10000000000000000000000000000000000  100000000000000000000000000000000000  0000000100000000000000000000000000 v           ~~~~~~ </pre>
---	---

  

**Вероятности:**

- q[1] = 0,58
- q[2] = 0,12
- q[3] = 0,06
- q[4] = 0,02
- q[5] = 0
- q[6] = 0
- q[7] = 0,02
- q[8] = 0,04
- q[9] = 0
- q[10] = 0,04
- q[11] = 0,02
- q[12] = 0,02
- q[13] = 0
- q[14] = 0
- q[15] = 0,02
- q[16] = 0
- q[17] = 0
- q[18] = 0,02
- q[19] = 0
- q[20] = 0
- q[21] = 0,02
- q[22] = 0
- q[23] = 0
- q[24] = 0
- q[25] = 0
- q[26] = 0
- q[27] = 0
- q[28] = 0,02
- q[29] = 0
- q[30] = 0
- q[31] = 0
- q[32] = 0

**Строка состояния:**

---



## Результати роботи програми з пробілами

```

327015 0.154086
о 152292 0.0931342
е 109122 0.0667336
а 105671 0.0646231
и 90629 0.0554242
н 84617 0.0517475
т 77729 0.0475352
с 70572 0.0431583
л 64887 0.0396817
в 60707 0.0371254
р 57550 0.0351947
к 43126 0.0263737
д 40011 0.0244687
м 38363 0.0234609
п 35247 0.0215553
у 34631 0.0211786
я 27973 0.0171069
ы 25537 0.0156172
ь 25196 0.0154086
г 25027 0.0153053
б 22386 0.0136902
з 22116 0.013525
ч 19176 0.0117271
й 14756 0.00902403
ж 14258 0.00871948
ш 11761 0.00719244
х 11162 0.00682612
ю 7735 0.00473034
ц 4734 0.00289508
э 3998 0.00244498
щ 3935 0.00240645
ф 2747 0.00167993
ъ 523 0.000319841
H1 : 4.28745
H2 : 3.953165
Всего символов в тексте : 1635189
Всего биграмм в тексте : 732620
```

## Результати роботи програми без пробілів

```

о 152292 0.116416
е 109122 0.0834155
а 105671 0.0807775
и 90629 0.069279
н 84617 0.0646833
т 77729 0.0594179
с 70572 0.053947
л 64887 0.0496012
в 60707 0.0464059
р 57550 0.0439926
к 43126 0.0329666
д 40011 0.0305854
м 38363 0.0293256
п 35247 0.0269437
у 34631 0.0264728
я 27973 0.0213832
ы 25537 0.0195211
ь 25196 0.0192604
г 25027 0.0191312
б 22386 0.0171124
з 22116 0.016906
ч 19176 0.0146586
й 14756 0.0112798
ж 14258 0.0108992
ш 11761 0.00899039
х 11162 0.0085325
ю 7735 0.00591282
ц 4734 0.00361878
э 3998 0.00305617
щ 3935 0.00300801
ф 2747 0.00209987
ъ 523 0.000399794
H1 : 4.45686
H2 : 3.99106
Всего символов в тексте : 1308174
Всего биграмм в тексте : 527085
```

bigrams неперекрестные.txt																																							
	а	б	в	г	д	е	ё	ж	з	и	й	к	л	м	н	о	п	р	с	т	у	ф	х	ц	ч	ш	щ	ъ	ы	ь	э	ю	я						
а	5	647	4143	676	3376	24	0	1040	3928	66	0	5601	5016	2051	8250	2	963	6291	1136	4153	7	85	553	445	1508	1115	295	0	0	0	0	0	0	1504					
б	449	3	14	0	24	858	0	0	38	145	314	1	0	19	19	6	2592	0	94	33	19	437	0	0	0	0	0	0	205	67	0	322	35	4994					
в	2476	78	26	18	810	1195	0	0	624	1801	0	341	3	1	92	39	66	3753	7	228	15	10	963	0	82	10	5	5	13	0	815	8	0	245	10640				
г	602	6	37	1	7	2574	0	0	7	271	370	0	1	92	39	66	3753	7	228	15	10	963	0	1	0	0	0	0	0	138	14	1	4	62	3006				
д	1376	15	258	846	22	2119	0	582	495	1102	112	0	190	0	473	3734	0	212	245	147	1265	0	1	0	0	0	0	0	0	93	36	2	404	407	5405				
е	780	1615	3395	392	3404	1279	0	2880	179	2310	1	292	3394	2121	6638	1883	1839	4371	2529	3267	180	96	25	675	2469	1828	961	103	964	1053	0	0	96	3156					
ё	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0					
ж	932	5	1	0	13	840	0	14	106	326	0	7	274	0	7	1368	0	215	23	0	1044	0	0	0	0	0	0	0	0	11	0	0	1	417	1464				
з	2952	0	402	1	1	937	0	0	29	1622	0	1	1	3	9	761	0	22	22	2	686	0	0	0	0	0	0	0	0	40	107	3	11	508	2546				
и	87	546	2424	516	1815	156	0	1127	273	750	0	1876	4994	2356	6345	631	569	3924	1010	2861	6	397	128	204	1204	1391	577	0	14	120	0	0	11	8408					
й	414	0	0	0	0	1	1827	0	0	0	920	0	0	0	0	0	2827	0	0	0	0	0	0	0	0	0	0	0	1020	0	6	0	32	21					
к	2949	114	105	55	120	765	0	42	102	1505	29	4	341	41	160	1235	61	177	3853	281	537	0	0	54	159	222	0	122	804	38	0	118	6743	0					
л	6673	548	822	1029	695	4684	0	1	149	3444	7	281	78	102	2	5212	587	55	1911	159	863	52	82	1	14	336	0	1870	0	54	36	657	1882	0					
м	2109	15	101	0	102	3037	0	5	219	2333	20	3	4	16	0	3720	0	283	699	6	740	1	38	0	4	9	0	0	1137	200	1	10	293	3853	0				
н	3266	230	770	184	1365	5938	0	972	1342	2240	337	851	215	697	2021	4886	61	458	836	736	203	0	68	0	390	305	35	0	146	864	10	11	431	11466					
о	4	1695	5965	6693	2845	504	0	20	560	102	10	6172	4686	2729	6984	170	6681	4918	2249	11594	3	53	1488	233	45	161	2	0	0	17	0	0	0	8855					
п	765	0	169	0	78	718	0	0	0	126	0	0	190	118	4	1057	28	56	1283	45	552	0	0	0	0	0	1	16	0	70	0	4	0	28	11051				
р	2413	792	558	815	1294	4961	0	3	189	340	2	1085	0	30	50	4963	4873	5	261	2013	420	437	69	0	19	1	1	0	226	0	0	36	49	2754					
с	2612	64	2439	72	330	3015	0	19	146	2234	623	77	799	72	405	5659	1	143	777	775	942	7	51	0	0	0	0	0	637	624	21	50	505	10983					
т	3586	6	156	2	141	2747	0	0	16	2459	130	369	22	3	274	5263	36	397	8827	86	1031	12	33	0	3049	46	0	0	576	46	1825	397	1090	5404					
у	89	722	617	472	1235	61	0	105	208	5	0	1124	881	1816	1536	40	518	2169	611	1052	2	66	79	448	429	199	71	0	2	0	0	0	0	2470					
ф	381	0	0	0	6	9	0	0	0	19	0	1	5	4	214	24	8	3	6	9	1	0	0	0	0	0	0	0	0	0	0	0	0	596					
х	696	105	32	0	36	699	0	0	0	1423	1	0	7	0	9	331	0	70	133	9	242	0	0	0	0	0	0	0	845	1	10	9	179	725					
ц	41	1	8	1	122	147	0	0	2	868	10	10	0	12	549	42	12	47	18	85	4	0	0	0	0	1	0	0	9	64	0	9	33	254					
ч	567	12	29	22	33	714	0	17	12	947	36	13	156	23	110	970	95	39	227	144	560	0	0	0	0	0	0	0	161	55	0	69	94	4068					
ш	954	3	97	4	85	496	0	6	252	94	8	7	6	21	780	0	240	73	0	500	0	10	0	73	0	0	0	0	346	393	21	13	502	0					
щ	201	172	1	0	0	410	0	0	0	122	0	0	3	83	113	1	26	1	10	223	0	0	0	0	0	0	0	0	0	2	4	0	368	179	43				
ъ	0	82	8	0	113	0	0	0	13	0	0	0	0	0	0	0	0	0	24	29	0	0	3	0	0	0	0	0	0	0	0	0	0	0	0				
ы	0	3270	1886	1	434	0	0	0	469	0	0	0	655	654	2523	0	227	1202	271	997	0	5	0	117	0	0	0	0	0	0	0	0	0	0	1				
ь	0	23	107	0	283	0	0	32	280	0	0	2	2938	40	702	0	726	771	2106	4268	0	3	0	0	126	150	10	0	0	0	0	0	0	0	3				
э	1	0	0	0	0	0	0	0	0	0	0	0	1	0	15	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1873					
ю	568	5	1	0	7	168	0	35	26	324	0	2	863	20	87	319	0	83	72	32	711	0	0	0	0	0	0	0	69	0	306	0	27	90	12				
я	1692	294	171	0	292	251	0	0	312	1599	0	0	1384	353	1692	506	247	796	2443	397	50	0	0	6	0	1	0	78	0	411	0	0	116	889	0				
	8931	122	3937	404	638	9919	0	92	793	11354	4029	2744	4087	3625	2318	12917	11	574	1744	2427	2614	67	1950	69	97	38	7	0	2278	5088	0	1228	5862	0	0				

bigrams перекрестные.txt																																							
	а	б	в	г	д	е	ё	ж	з	и	й	к	л	м	н	о	п	р	с	т	у	ф	х	ц	ч	ш	щ	ъ	ы	ь	э	ю	я						
а	8	1281	8114	1384	6704	59	0	2128	7742	135	2	0	11145	10138	4101	16381	5	1887	12629	2238	8281	20	172	1141	884	2983	2213	598	0	0	0	0	0	0	2958				
б	907	7	21	0	43	1696	0	77	304	628	0	0	33	41	17	5313	0	204	72	38	869	0	0	0	0	0	0	0	385	122	0	621	68	10001					
в	5004	147	53	42	1615	2447	0	0	1262	3667	1	723	4	1	47	11456	0	436	2746	4221	1307	0	173	32	9	21	1	0	1694	16	1	0	491	21414	0				
г	1169	15	65	3	10	5201	0	11	534	732	0	1	156	78	123	7427	14	413	31	24	1890	0	2	0	0	0	0	0	108	29	4	6	117	6043	0				
д	2778	40	477	1709	39	4131	0	1209	974	2222	225	0	416	0	911	7443	0	429	449	258	2522	0	2	0	0	0	0	0	166	73	3	788	834	10893					
е	1576	3260	6760	819	6830	2557	0	5757	365	4582	2	583	6700	4325	13499	3818	3712	9011	5141	6513	342	177	45	1364	4921	3633	1903	214	1848	2080	0	0	192	6198	0				
ё	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0				
ж	1844	10	1	1	26	1648	0	34	207	643	0	20	551	0	15	2734	0	447	36	0	1981	0	0	0	0	0	0	0	24	0	0	7	849	2964	0				
з	5947	0	750	1	2	1850	0	0	59	3269	6	4	19	1532	0	43	42	7	1407	0	0	0	0	0	0	0	0	0	81	209	5	26	1026	5199	0				
и	172	1084	4822	1042	3813	292	0	2216	553	1536	0	3711	9887	4756	12739	1291	1130	7778	2168	5583	14	828	241	391	2354	2830	1172	0	29	217	0	0	21	17059	0				
й	796	0	0	0	0	1	3747	0	0	0	1813	0	0	0	0	0	0	5634	0	0	0	109	0	0	0	0	0	0	2062	0	15	0	65	43	0				
к	5975	244	201	108	27	1563	0	107	194	3009	62	15	684	89	288	2440	111	347	7806	570	1406	0	0	108	325	472	0	0	250	1674	74	6	250	13455	0				
л	13313	1137	1708	2066	1332	9227	0	1	344	6822	56	582	144	190	52	10434	1181	320	3954	214	1650	104	157	2	34	557	0	0	3755	93	106	8	1340	3743	0				
м	4193	46	439	1	21	6057	0	10	456	4571	6	6	7	29	0	7336	1	560	1435	17	1450	1	83	3	10	13	0	0	225	393	3	15	564	355	0				
н	6714	446	1613	366	2668	11889	0	1954	2553	4631	617	1707	465	1458	4015	9940	123	926	1660	1502	417	0	144	0	766	585	55	0	303	1694	27	23	845	22531	0				
о	9	3332	11838	13516	5168	999	0	0	53	1083	223	14	12357	9277	5328	13962	325	13397	9978	5040	23473	6	110	3038	467	97	341	2	0	1	36	0	0	0	17324	0			
п	1588	0	349	0	170	1437	0	0	0	253	0	0	373	283	8	2114	4	116	2561	96	1094	0	0	0	0	0	0	0	141	0	13	4	47	21888	0				
р	9252	124	1475	155	20	1689	0	37	349	6	2198	0	3	279	4447	1205	169	1660	164	769	1279	8	336	493	1508	1963	173	1353	874	13	108	0	0	1027	22048	0			
с	7212	11	299	4	310	5541	0	3	41	3811	303	730	52	9	571	10430	66	810	17533	150	2113	25	54	0	5997	115	1	0	1191	91	3674	769	2230	10885	0				
т	1877	1438	1242	937	2546	118	0	206	404	7	0	2195	1706	3636	3152	89	1006	4309	1251	2063	4	133	173	896	916	426	154	0	5	0	0	0	0	4982	0				
у	744	0	0	0	0	1	0	0	0	38	0	0	7	15	9	405	46	18	12	9	16	7	0	0	0	0	0	0	0	0	0	3	0	0	1149	0			
ф	1449	186	61	0	9	1400	0	0	2846	2	0	13	0	16	668	0	129	259	15	468	0	0	0	0	0	0	0	0	1636	1	14	13	323	1447	0				
х	966	1	20	3	242	300	0	0	5	1704	26	18	1	36	1088	79	31	84	40	147	15	0	0	0	0	0	1	0	14	121	0	19	59	526	0				
ц	1160	29	59	43	60	1381	0	51	25	1987	80	27	291	39	215	1898	205	90	457	282	1160	2	1	0	0	0	0	0	334	103	0	125	197	8230	0				
ч	1968	6	1869	9	194	964	0	0	17	514	182	15	10	12	26	1516	2	461	133	2	1005	0	22	0	147	1	0	0	711	783	0	49	30	1018	0				
ш	383	310	2	0	0	858	0	0	0	241	0	0	4	7	164	234	5	50	1	18	465	0	0	0	0	0	0	0	3	11	0	727	358	76	0				
щ	1659	17	0	0	0	1	0	25	0	0	0	0	0	0	0	0	0	38	5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0			
ъ	0	6475	3706	1	894	0	0	0	910	0	0	0	1376	1390	5038	0	466	2459	504	2041	0	12	0	246	0	0	0	0	0	0	0	0	0	0	1	0			
ы	0	55	242	0	569	0	0	62	545	0	0	2	5854	69	1442	0	1452	1571	4200	8524	0	8	0	0	255	317	21	0	0	0	0	0	0	0	3	0			
ь	0	5	1	0	0	0	0	0	63	596	0	0	0	1	0	30	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3693	0			
ю	1161	5	5	1	0	15	347	0	60	38	596	0	5	1766	39	198	598	0	164	174	63	142	0	6	0	0	0	150	0	637	0	52	186	21	0				
я	619	35	0	0	53	92	0	667	3223	0	68	3263	988	42	1574	4759	8	49	15	474	869	209	7	2	0	0	0	0	0	0	0	0	0	2	1760	0			
	18083	250	7945	812	1273	19741	0	169	1697	22510	8042	5384	8251	7337	4588	26099	24	1161	351	1468	5201	150	3945	144	193	70	19	19	0	4531	10156	0	2338	11803	0				

## **Висновки**

Під час виконання даного практикуму ми навчилися рахувати ентропію на символ джерела та знаходити надлишковість; навчилися порівнювати різні моделі джерела відкритого тексту для наближеного визначення ентропії.

По результатам нашої програми ми можемо бачити, що дані, які ми отримали є близькими до середньостатистичних даних, які можна знайти в інтернеті.