

Національний технічний університет України
«Київський політехнічний інститут імені Ігоря Сікорського»
Фізико-технічний інститут

# ЛАБОРАТОРНА РОБОТА №1

з дисципліни

«Криптографія»

на тему: «Експериментальна оцінка ентропії на символ джерела відкритого тексту»

Виконали:

студенти 3 курсу ФТІ

групи ФБ-74

Пудім Єлизавета та Горобець Ангеліна

Перевірили:

Чорний О.

Савчук М. М.

Завадська Л. О.

## 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  $1 M \delta$ ), 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  $H^{10}$ ,  $H^{20}$ ,  $H^{30}$ .
- 3. Використовуючи отримані значення ентропії, оцінити надлишковість російської мови в різних моделях джерела.

# Результати виконнання програми:

## **Space:**

```
' ': 0.1672034781394746, 'o': 0.09603150069552818,
'e': 0.07314077091445752, 'a': 0.06929292954474066,
'h': 0.05774691250305384, 'n': 0.05498225032034183,
't': 0.04991661140666211, 'c': 0.04395787941186736,
'n': 0.042096905273549486, 'b': 0.03919887120014758,
'p': 0.03280948510971396, 'k': 0.028577108896279048,
'д': 0.024793460539370885, 'м': 0.02389974422512178,
'y': 0.0227230800680072, 'n': 0.020107370604336705,
'я': 0.018093081115038864, 'ь': 0.016695169195331237,
'ы': 0.0154430938289948, 'г': 0.015148304556582089,
'6': 0.014661559478877383, '4': 0.014224049818763804,
'3': 0.013658777366166917, 'ж': 0.009569121540432872,
'й': 0.00857070205966086, 'ш': 0.007248824582309153,
'х': 0.006379414360288582, 'ю': 0.0052457532894244815,
'э': 0.0029697058838193723, 'щ': 0.002367039443178589,
'u': 0.002161995742071228, 'ф': 0.0008388717984513903,
```

# 'ъ': 0.00024617708795564574})

Monograms
4.348869152023138
Entropy bigrams with step1
3.939835313253219
Entropy bigrams with step2
3.9398911732225845

Redundancy

Monograms – 0.132

Entropy bigrams with step1 – 0.214

Entropy bigrams with step2 – 0.213

■ 0.0055 0.0000 # 0.0042 0.0000 © 0.0000 0.0013 0.0024 0.0023 0.0000 x 0.0014 0.0000 **5** 0.0055 0.0000 0.000 # 0.0001 0.0004 0.0005 0.0002 0.0003 0.0000 X 0.0000 x 0.0078 0.0000 # 0.0077 0.0000 0.0001 0.0000 0.00052 0.0000 M 0.0024 0.0000 # 0.1019 6 .0000 0.0034 0.0001 0.0024 0.0000 0.0019 0.0010 0.0000 0.0010 0.0000 0.0010 0.0000 0.0010 0.0000 0. # 0.0012 0.0000 p 0.0067 0.0002 0.0004 0.00050 0.0000 0.0004 0.0000 0.0004 0.0000 0.0004 0.0000 0.0004 0.0000 0.0004 0.0000 - 0.018 0.0000 0.0018 0.0000 0.000 5 0.0009 0.0000  $\pm$  0.0008 0.0000 0.0002 0.0000 0.000 € 0.0000

#### Without spaces

'o': 0.11531208185282417, 'e': 0.08782549998054259,

'a': 0.08320511400544209, 'h': 0.06934096263279681,

'и': 0.06602122952856197, 'т': 0.05993854452603251,

'c': 0.05278345701259942, 'л': 0.05054884857047922,

'B': 0.04706896603335299, 'p': 0.03939676049128458,

'к': 0.0343146352634084, 'д': 0.02977133055740983,

'm': 0.028698179684670255, 'y': 0.027285272538413414,

'n': 0.02414439791296851, 'я': 0.021725692459206795,

'ь': 0.02004711686119085, 'ы': 0.018543657932785132,

'r': 0.018189682784384983, '6': 0.017605212190514963,

'4': 0.017079862181684294, '3': 0.016401097996485692,

'ж': 0.011490347629040032, 'й': 0.01029147196786235,

'ш': 0.008704196513831223, 'x': 0.007660231752693355,

'ю': 0.006298961573116448, 'э': 0.00356594414825946,

'щ': 0.0028422782529043925, 'ц': 0.002596067208879762,

'ф': 0.0010072950311767542, 'ъ': 0.0002956029251967443

Entropy in monogram 4.439982265826084 Entropy bigrams with step1 4.11963646406772 Entropy bigrams with step2

4.118897060849513

#### Redundancy

Monograms - 0.112

Entropy bigrams with step1 - 0.17625

Entropy bigrams with step2 - 0.17623

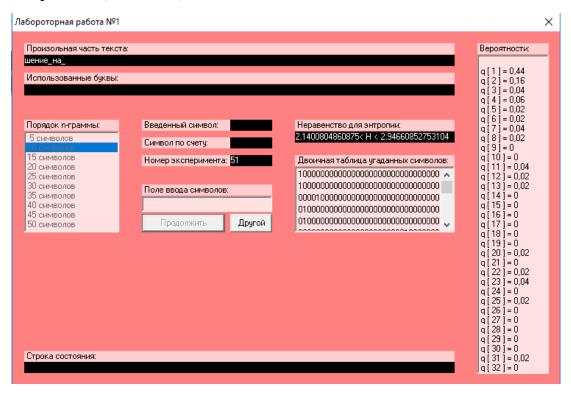
# **Bigrams table**

A	0	H	n	**	M	**	*		p,	B		×	Ħ	H	e	n	•	
0.0035 0.	0020	0.0000 0.	a 0.0005 0.0015 0.0057 0.0012 0.0035 0.0020 0.0000 0.0014 0.0056 0.001	Ψ	0.0006 0.0063 0.0122	0.0122 0.0	0.0037 0.0075	0.0020	0.0024 0.0031	0.0064 0.0059	0.0006	0.0003 0.0012	0.0002	0.0016 0.0009 0	0.0003 0.0000	00000 0.0000 01	00 0.0004 0.0010	0 0.0032
.0000 0	0024	0.0000 0.	0.0000 0.0001 0.0000 0.0000 0.0024 0.0000 0.0000 0.0000 0.00	9	0.0000 0.0003	0.000	0.0001 0.0004	0.0024	0.0000 0.0013	0.0001	0.0000 0.0012 0.0000	.0000 0.0001	0.000 0.000 1	0.0000	0.0002 0.0001	0.0050 0.0000	00 0.0001 0.0000	9000.000
0 8000"	.0053	0.0000 0.	0.0066 0.0002 0.0003 0.0003 0.0008 0.0053 0.0000 0.0001 0.0008 0.0056	8 0.0056 0.0	0.0000 0.0007	0.0010	0.0004 0.0020	0.0089	0.0009 0.0014	0.0049	0.0007 0.0009 0.0000	.0000 0.0001	0.0000 0.0003	0.0013	0.0000 0.0000	0 0.0029 0.0002	02 0.0003 0.0000	0 0.0003
.0015	0000	0.0000 0.0001 0.0000 0.0015 0.0004 0.0000 0.0000	.0000 0.000	0.0000 0.0010 0.0	0.0000 0.0002	0.0018	0.0000 0.0003	0.0100	0.0001 0.0008	0.0001	0.0000 0.0008 0.0000	0000 0.0000	0.0000 0.0001	0.000	0.0000 0.0000	0000.0000.0000	00 0.0000 0.0000	0000.000
10001	0.0055	0.0001 0.0010 0.0000 0.0001 0.0055 0.0000 0.0000	0000 0.000	0.0000 0.0026 0.0	0.0000 0.0003	0.0010 0.0	0.0001 0.0020	0.0045	0.0001 0.0018	0.0005	0.0003 0.0022	0.0000 0.0001	0.0002 0.0001	0.0001	0.0000 0.0001	00000 0.0000	0000.00000.0000	0 0.0005
0.0040	0.0031	e 0.0003 0.0028 0.0050 0.0049 0.0040 0.0031 0.0000 0.0012	.0012 0.002	0.0022 0.0017 0.0	0.0031 0.0030	0.0074 0.0	0.0062 0.0103	0.0019	0.0038 0.0068	0.0066	0.0064 0.0007	0.0001 0.0012	0.0002 0.0017	0.0012	0.0008 0.0000	00 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	00 0.0004 0.0005	9000.009
0000	0.000	0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000	.0000 0.000	0.0000 0.0000 0.0	0.0000 0.0000	0.000.0	0.0000 0.0000	0.000.0	0.0000 0.0000	0.0000	0.0000 0.0000	0.0000 0.0000	0.0000 0.0000	0.000.0	0.0000 0.0000	00 00 00 00 00 00 00 00 00 00 00 00 00	00 0.0000 0.0000	0000.000
0000	0.0050	0.0000 0.	0.0017 0.0001 0.0000 0.0000 0.0009 0.0050 0.0000 0.0000 0.0000	00	0.0000 0.0001	0.000.0	0.0000 0.0012	0.0001	0.0000 0.0000	0.0001	0.0000 0.0003	0.0000 0.0000	0.0000 0.0001	0.000	0.0000 0.0000	10 0.0000 0.0001	01 0.0000 0.0000	0000.000
.0010	0.0003	0.0010 0.0007 0.0010 0.0003 0.0000 0.0001	.0001 0.000	0.0001 0.0004 0.0	0.0000 0.0002	0.0002	0.0004 0.0023	0.0005	0.0001 0.0003	0.0002	0.0001 0.0003	0.0000 0.0000	0.0000 0.0001	0.000.0	0.0000 0.0000	0 0.0004 0.0002	02 0.0000 0.0000	9000.000
.0027	0.0034	0.0000 0.	0.0004 0.0014 0.0054 0.0009 0.0027 0.0034 0.0000 0.0006 0.0026 0.001	o,	0.0015 0.0029	0.0057	0.0040 0.0070	0.0020	0.0023 0.0011	0.0046 0.0056	1056 0.0006 0.0001	.0001 0.0016	0.0016 0.0010 0.0029	0.000.0	0.0002 0.0000	00 0.0000 0.0000	00 0.0003 0.0003	13 0.0023
.0007	0.0001	0.0000 0.	0.0002 0.0007 0.0002 0.0007 0.0001 0.0000 0.0002 0.0002 0.0008		0.0000 0.0006	0.0003 0.6	0.0003 0.0003 0.0009	0.0005	0.0007 0.0003	0.0013	0.0006 0.0002 0.0000	.0000 0.0001	0.0001	0.0003 0.0001 0	0.0000 0.0000 0.0000	00 0 0 0000 0 0 0000	00 0.0001 0.0000	1000.001
.0002	0.0007	0.0000 0.	0.0095 0.0004 0.0006 0.0001 0.0002 0.0007 0.0000 0.0002 0.0001 0.0036	1 0.0036 0.0	0.0000 0.0003	0.0005	0.0002 0.0014	0.0097	0.0004 0.0018	0.0017	0.0007 0.0018	0.0000 0.0001	0.0000 0.0002	0.000	0.0000 0.0000	00 0 0 0000 0 0000	00 0.0001 0.0000	1000.0001
.0003	0.0069	0.0000 0.	0.0054 0.0002 0.0009 0.0003 0.0003 0.0069 0.0000 0.0005 0.0002 0.00	88	0.0000 0.000.0	0.0007	0.0002 0.0010	0.0081	0.0006 0.0001	0.0023	0.0004 0.0016 0.0000	0000 0.0000	0.0000 0.0007	0.000	0.0000 0.0000	0 0.0012 0.0042	42 0.0002 0.0012	2 0.0022
.0004	0.0036	0.0003 0.0007 0.0003 0.0004 0.0036 0.0000 0.0001	.0001 0.000	0.0002 0.0038 0.0	0.0000 0.0005	0.0003	0.0003 0.0022	0.0045	0.0008 0.0003	6000.0	0.0004 0.0035	0.0000 0.0001	0.0000 0.0005	0.0001	0.0000 0.0000	10 0.0010 0.0001	01 0.0001 0.0000	9000.0 0
0100	0.0120	0.0000 0.	0.0003 0.0007 0.0002 0.0010 0.0120 0.0000 0.0001 0.0003 0.0094		0.0000 0.0005	0.0001	0.0001 0.0042	0.0118	0.0007 0.0002	0.0019	0.0008 0.0033	0.0000 0.0001	0.0003 0.0005	0.000	0.0002 0.0000	0 0.0033 0.0015	15 0.0001 0.0002	2 0.0022
.0062	0.0039	0.0055 0.0114 0.0060 0.0062 0.0039 0.0000 0.0029	.0029 0.001	0.0017 0.0024 0.0	0.0035 0.0031	0.0070 0.0	0.0065 0.0130	0.0030	0.0036 0.0071	0.0095	0.0095 0.0008	0.0002 0.0006	5 0.0001 0.0034	0.0016	0.0002 0.0000	00 00 00 00 00 00 00 00 00 00 00 00 00	00 0.0005 0.0010	.0 0.0011
0000	0.0024	0.0000 0.0000 0.0000 0.0000 0.0024 0.0000 0.0000	.0000 0.000	0.0000 0.0009 0.0	0.0000 0.0001	0.0005	0.0000 0.0001	0.0098	0.0000 0.0073	0.0073 0.0000 0.0	0.0000 0.0006 0.0000	0000 0.0000	0.0000 0.0000 0.0000	0.000.0	0.0000 0.0000	0 0.0003 0.0000	00 0.0000 0.0000	0 0.0004
.0004	0.0003 0.0004 0.0004 0.0004 0.0062	0.0000 0.	0.0000 0.0004 0.0001 0.0053		0.0000 0.0007	0.0001	0.0001 0.0008	0.0081	0.0002 0.0000	0.0003	0.0006 0.0027	0.0000 0.0001	1 0.0000 0.0001	0.0003	0.0000 0.0000	0 0.0014 0.0010	10 0.0000 0.0001	11 0.0011
000	5 0.0054	0.0000 0.	0.0021 0.0002 0.0024 0.0001 0.0006 0.0054 0.0000 0.0002 0.0001 0.001	·	0.0000 0.0061	0.0030	0.0011 0.0016 0.0030	0.0030 0.0	0.0022 0.0002	0.000	0.0125 0.0007	0.0000 0.0002	0.0002 0.0001 0.0006	0.0001	0.0000 0.000	0.0000 0.0004 0.0035	35 0.0001 0.0001	0.0036
000	1 0.0061	0.0000 0.	0.0003 0.0035 0.0001 0.0004 0.0061 0.0000 0.0001 0.0001 0.0048		0.0000 0.0008	0.0004 0.0	0.0004 0.0002 0.0017	0.0175	0.0004 0.0031	0.0015	0.0004 0.0015 0.0000	.0000 0.0001	0.0001 0.0005	0.000.0	0.0000 0.0000	0 0.0018 0.0076	76 0.0001 0.0000	10 0.0007
.0022	0.0004	0.0000 0.	0.0002 0.0009 0.0020 0.0012 0.0022 0.0004 0.0000 0.0022 0.0005 0.0008		0.0001 0.0014	0.0015	0.0017 0.0012	0.0005	0.0012 0.0007	0.0019	0.0017 0.0002 0.0000	.0000 0.0004	1 0.0000 0.0015	0.000	0.0002 0.000	0.0000 0.0000 0.0000	00 0.0001 0.0011	1 0.0003
0000	0.0001	0.0001 0.0000 0.0000 0.0000 0.0000 0.0001 0.0000 0.0000	.0000 0.000	0.0000 0.0004 0.0	0.0000 0.0000	0.000	0.0000 0.0000	0.0001	0.0000 0.0002	0.000	0.0000 0.0000 0.0000	0.0000 0.0000	0.0000 0.0000	0.000	0.0000 0.0000	10 0.0000 0.0001	01 0.0000 0.0000	0000.000
.0002	0.0001	0.0000.0	0.0010 0.0001 0.0004 0.0001 0.0002 0.0001 0.0000 0.0000 0.0001 0.0004		0.0000 0.0002	0.0002	0.0002 0.0004	0.0029	0.0003 0.0002	0.0004	0.0001 0.0002 0.0000	0.0000 0.0000	0.0000 0.0001	0.000	0.0000 0.0000	00.0000 0.0000	00 0.0000 0.0000	0000.000
0000	0.0007	0.0000 0.	щ 0.0006 0.0000 0.0001 0.0000 0.0000 0.0007 0.0000 0.0000 0.0000 0.0002	0 0.0002 0.0	0000 0.0001	0.0000 0.0	0000 0.0000	0.0004 0.0	0000 0.0000	0.00000.0	000 0.0003	0000 0.0000	0.0000 0.0	0 000 0.0000 0	.0000 0.000	0 0.0002 0.00	0.0000 0.0001 0.0000 0.0	0000.0

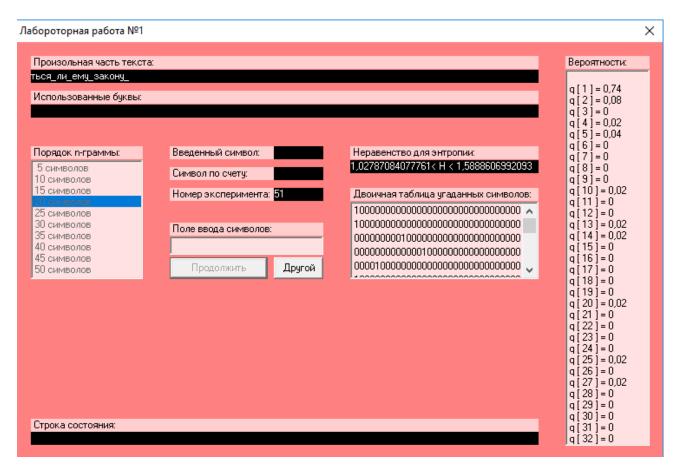
■ 0.001 0.0000 0.0001 0.0000

# Скріншоти результатів рожевої программи:

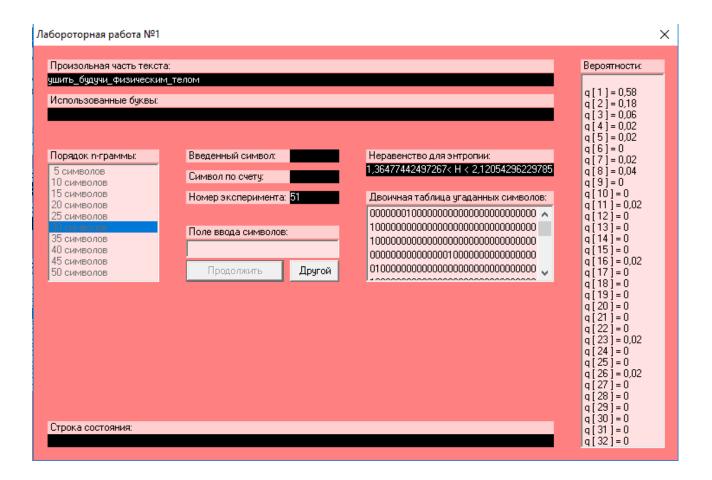
# 10-грамми: 0,57 < R < 0,41



20-грамми 0.79 < R < 0.68



30-грами: 0,72 < R < 0,57



# Висновки:

Під час данного комп'ютерного практикуму, ми навчились визначати ентропію на символ джерела та його надлишковості. Порівняли різні моделі джерел відркритого тексту для наближенного визначення ентропії та набули практичних навичок оцінки ентропії на символ джерела.