МІНІСТЕРСТВО ОСВІТИ І НАУКИ, МОЛОДІ ТА СПОРТУ УКРАЇНИ НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ УКРАЇНИ «КИЇВСЬКИЙ ПОЛІТЕХНІЧНИЙ ІНСТИТУТ» ФІЗИКО-ТЕХНІЧНИЙ ІНСТИТУТ

Лабораторна робота №1 3 дисципліни «Криптографія»

Виконали:

Пасько Олександр ФБ-84

Завгородня Анастасія ФБ-81

Перевірив:

Чорний О. М.

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.

Хід роботи:

Для виконання цієї лабораторної роботи, в якості експериментального тексту, було використано «Экспансия. Пробуждения Левиафана». Перш за все, була створена програма, яка очищала текст від непотрібних символів (цифри, знаки пунктуації, тощо). Лише після цього була створена програма для обробки тексту та виконання завдання. Так як по завданню необхідно працювати як з пробілом, так і без нього, то робота програми відбувається в два етапи, які відрізняються один від одного лише алфавітом.

Що стосується труднощів, то найбільше їх виникло при роботі с біграмами, де потрібно було враховувати перетин (або його відсутність) букв, який впливав на підрахунок частот та ентропії, та вивід результатів в таблиці

Результати:

1. Написати програми для підрахунку частот букв і частот біграм в тексті, а також підрахунку Н1 та Н2 за безпосереднім означенням. Підрахувати частоти букв та біграм, а також значення Н1 та Н2 на довільно обраному тексті російською мовою достатньої довжини (щонайменше 1Мб), де імовірності замінити відповідними частотами. Також одержати значення Н1 та Н2 на тому ж тексті, в якому вилучено всі пробіли.

Results with spaces:

• H letter: 4.389386586883138

• H cross bigrams: 3.999922042624518

• H non-cross bigrams: 3.9971312787557705

Results without spaces:

• H letter: 4.457568779065807

• H cross bigrams: 4.148077673506048

• H non-cross bigrams: 4.147670599351096

Results with spaces	Results without spaces
"":0.154650	o:0.108470
o:0.091700	a:0.083020
a:0.070180	e:0.082070
e:0.069380	и:0.067530
и:0.057080	н:0.065400
н:0.055280	л:0.059630
л:0.050410	т:0.059300
т:0.050130	c:0.053750
c:0.045440	p:0.046810
p:0.039570	в:0.040960
в:0.034620	к:0.034430
к:0.029110	м:0.032930
м:0.027840	д:0.031430
д:0.026570	п:0.029310
п:0.024780	y:0.028330
y:0.023950	я:0.019890
я:0.016810	ы:0.018540
ы:0.015670	з:0.018330
3:0.015500	ь:0.018030
ь:0.015240	б:0.016180
6:0.013670	г:0.015080
г:0.012740	ч:0.014060
ч:0.011880	x:0.010530
x:0.008910	й:0.009970
й:0.008420	ж:0.009400
ж:0.007950	ш:0.008040
ш:0.006800	ю:0.005510
ю:0.004660	ц:0.003750
ц:0.003170	э:0.003740
э:0.003160	щ:0.003380
щ:0.002850	ф:0.001970
ф:0.001660	ъ:0.000230
ъ:0.000190	ë:0.000010
ë:0.000010	

Results with spaces (cross)

о п $\pi = 0.00390 \ 0.0004 \ 0.00099 \ 0.0001 \ 0.00017 \ 0.00061 \ 0.00000 \ 0.00001 \ 0.00017 \ 0.00061 \ 0.00000 \ 0.00001 \ 0.$ x 0.00096 0.00007 0.00000 0.00001 0.00078 0.00286 0.00000 0.00001 0 $3 - 0.00582 \ 0.00014 \ 0.00104 \ 0.00014 \ 0.00104 \ 0.00001 \ 0.00007 \$ $\mathbf{H} = 0.00035\ 0.00044\ 0.00225\ 0.00063\ 0.00156\ 0.00188\ 0.00000\ 0.00031\ 0.00190\ 0.00003\ 0.00105\ 0.00188\ 0.00003\ 0.00012\ 0.00070\ 0.00085\ 0.00003\ 0.00012\ 0.00170\ 0.00060\ 0.00086\ 0.00054\ 0.00019\ 0.000000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.000000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.000000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.000000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.000000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.000000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.000000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.000000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.000000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.000000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.000000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.000000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.000000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.000000\ 0.00000\ 0.00000\ 0.00000\ 0.000000\ 0.00000\ 0.00000\ 0.000000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.000$ $\ddot{\mathbf{u}} = 0.00000 \ 0.00001 \ 0.00$ $0.00686\ 0.00002\ 0.00012\ 0.00000\ 0.00006\ 0.00000\ 0.000001\ 0.000000\ 0.000000\ 0.000000\ 0.000000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.000$ $\pi = 0.00594\ 0.00001\ 0.00002\ 0.00010\ 0.00002\ 0.00010\ 0.00001\ 0.00000$ $0.00005\ 0.00349\ 0.00646\ 0.00380\ 0.00378\ 0.00131\ 0.000000\ 0.00181\ 0.00166\ 0.00069\ 0.0035\ 0.00232\ 0.00736\ 0.00560\ 0.00656\ 0.00035\ 0.00155\ 0.00558\ 0.00703\ 0.00667\ 0.00013\ 0.00000\ 0.00014\ 0.00014\ 0.00072\ 0.00023\ 0.000000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.000000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.000000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.000000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.000000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.000000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.000000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.000000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.000000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.000000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.000000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.000000\ 0.00000\ 0.00000\ 0.00000\ 0.000000\ 0.00000\ 0.00000\ 0.000000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.000000\ 0.00000\ 0.00$ $p = 0.00809 \ 0.00015 \ 0.00054 \ 0.00016 \ 0.00054 \ 0.000016 \ 0.00001 \$ $\texttt{T} = 0.00599 \ 0.00004 \ 0.00207 \ 0.00001 \ 0.00001 \ 0.00001 \ 0.00001 \ 0.00001 \ 0.00000 \ 0.000001 \ 0.00001 \ 0.000001 \ 0.000001 \ 0.000001 \ 0.$ $v = 0.00004 \ 0.00088 \ 0.00064 \ 0.00081 \ 0.00146 \ 0.00016 \ 0.00001 \ 0.00016 \ 0.00001 \$ $\pi = 0.00053 \ 0.00000 \ 0.00004 \ 0.000000 \ 0.00000 \ 0.00000 \ 0.00000 \ 0.000000 \ 0.00000 \ 0.00000 \ 0.00000 \ 0.00000 \ 0.00000 \ 0.00000$

Results with spaces (non-cross)

б в г д е ё ж з и й к л м н $r = 0.00099 \ 0.00000 \ 0.00001 \ 0.0000001 \ 0.0000001 \ 0.000001 \ 0.000001 \ 0.000001 \ 0.000001 \ 0.000001 \ 0.000001 \ 0.000001 \ 0.000$ ${\tt e} \quad 0.00013 \ 0.00108 \ 0.00118 \ 0.00251 \ 0.00302 \ 0.00121 \ 0.00000 \ 0.000074 \ 0.00123 \ 0.000070 \ 0.00180 \ 0.00175 \ 0.00033 \ 0.00068 \ 0.0000155 \ 0.00008 \ 0.00003 \ 0.00008 \ 0.0000180 \ 0.00001 \ 0.00001 \ 0.00001 \ 0.00001 \ 0.00015 \ 0.00001 \$ x 0.00096 0.00006 0.00000 0.00001 0.00001 0.00001 0.00001 0.00001 0.00001 0.00001 0.00001 0.00001 0.00001 0.00001 0.00000 0 $\ddot{\text{w}} = 0.000000 \ 0.000001 \ 0.000001 \ 0.000001 \ 0.000001 \ 0.000002 \ 0.000001 \ 0.000002 \ 0.000001 \ 0.0000001 \ 0.0000001 \ 0.0000001 \ 0.0000001 \ 0.0000$ ${\tt H} = 0.01173\ 0.00001\ 0.00002\ 0.00010\ 0.00005\ 0.00084\ 0.00000\ 0.00005\ 0.00084\ 0.00000\ 0.00005\ 0.00008\ 0.00005\ 0.00008\ 0.00005\ 0.$ $0.00815\ 0.00012\ 0.00052\ 0.00015\ 0.00022\ 0.00043\ 0.00000\ 0.00038\ 0.00004\ 0.00003\ 0$ ${\tt c} \quad 0.00144 \ 0.00008 \ 0.00119 \ 0.00004 \ 0.00025 \ 0.00300 \ 0.00000 \ 0.00003 \ 0.00001 \ 0.00003 \ 0.00001 \ 0.00003 \ 0.00001 \ 0.00003 \ 0.00001 \ 0.00003 \ 0.00001 \ 0.00003 \ 0.00001 \ 0.00003 \ 0.00001 \ 0.00003 \ 0.00001 \ 0.00003 \ 0.00001 \ 0.00003 \ 0.00001 \ 0.00003 \ 0.00001 \ 0.00003 \ 0.00001 \ 0.00003 \ 0.00001 \ 0.00003 \ 0.00001 \ 0.00003 \ 0.00001 \ 0.00003 \ 0.00001 \ 0.00003 \ 0.00001 \ 0.00003 \ 0.00001 \ 0.000001 \ 0.00003 \ 0.00001 \ 0.00003 \ 0.00001 \ 0.00003 \ 0.00001 \ 0.00003 \ 0.00001 \ 0.00003 \ 0.00001 \ 0.00003 \ 0.00001 \ 0.000001 \ 0.00003 \ 0.00001 \ 0.00003 \ 0.00001 \ 0.00003 \ 0.000001 \ 0.00003 \ 0.00001 \ 0.00003 \ 0.00001 \ 0.00003 \ 0.00001 \ 0.00003 \ 0.00001 \ 0.00003 \ 0.00001 \ 0.00003 \ 0.00001 \ 0.00003 \ 0.00001 \ 0.00003 \ 0.00001 \ 0.00003 \ 0.00001 \ 0.000001 \ 0.00003 \ 0.00001 \ 0.00003 \ 0.00001 \ 0.00003 \ 0.00001 \ 0.00003 \ 0.00001 \ 0.00003 \ 0.00001 \ 0.00003 \ 0.00001 \ 0.00003 \ 0.00001 \ 0.00003 \ 0.000001 \ 0.0000001 \ 0.000001 \ 0.000001 \ 0.000001 \ 0.000001 \ 0.000$ $= \underbrace{0.00600\ 0.00004\ 0.00211\ 0.00001\ 0.00001\ 0.00001\ 0.00001\ 0.00001\ 0.000001\ 0.000001\ 0.000001\ 0.000001\ 0.000001\ 0.000001\ 0.000001\ 0.000001\ 0.0000$ $y = 0.00005 \ 0.00086 \ 0.00061 \ 0.00087 \ 0.00144 \ 0.00016 \ 0.00007 \ 0.00144 \ 0.00016 \ 0.00007 \ 0.000144 \ 0.00001 \ 0.00007 \ 0.000144 \ 0.00007 \ 0.00015 \ 0.00008 \ 0.00015 \ 0.00001 \ 0.00001 \ 0.00001 \ 0.00001 \ 0.00001 \ 0.00001 \ 0.000001 \ 0.00001$ $\Phi = 0.00025 \ 0.00000 \$ $= 0.00057\, 0.00000\, 0.00004\, 0.00000$ $= 0.00165\,0.00000\,0$ $_{
m II}$ 0.00038 0.00000 0.0000 0.0000 0.0000 0.0000 0.0000 0.00000 0.00000 0.00000 0.00000 0.00000 0.00000

ы	$0.00001\ 0.00033\ 0.00092\ 0.00017\ 0.00011\ 0.00095\ 0.00000\ 0.00006\ 0.00001\ 0.00003\ 0.00006\ 0.00001\ 0.00003\ 0.00006\ 0.00001\ 0.00003\ 0.000000\ 0.000000\ 0.000000\ 0.00000\ 0.00000\ 0.000000\ 0.00000\ 0.000000\ 0.00000\ 0.000$
Ь	$0.00000\ 0.00004\ 0.00001\ 0.00004\ 0.00001\ 0.00004\ 0.00007\ 0.00002\ 0.000000\ 0.00000\ $
э	$0.00000\ 0.000018\ 0.00000\ 0.000018\ 0.00000\ 0.00000\ 0.000000\ 0.000000\ 0.00000\ 0.000000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.000$
ю	$0.00000\ 0.00018\ 0.00000\ 0.000018\ 0.00000\ 0.00000\ 0.0000000\ 0.0000000\ 0.000000\ 0.000000\ 0.000000\ 0.000000\ 0.000000\ 0.000000\ 0.000000\ 0.000000\ 0.0000000\ 0.0000000\ 0.0000000\ 0.00000000$
я	$0.00000\ 0.00002\ 0.00002\ 0.00002\ 0.00002\ 0.00000\ 0.000015\ 0.00000\ 0.00001\ $
٠٠,	$^{\circ}$ 0.00262 0.00493 0.01298 0.00284 0.00667 0.00361 0.00000 0.00122 0.00518 0.00792 0.00000 0.00908 0.00213 0.00745 0.01617 0.01098 0.01686 0.00408 0.01488 0.00659 0.00400 0.00086 0.00283 0.00044 0.00512 0.00118 0.00011 0.00000 0.00000 0.00025 0.00003 0.00150 0.00000

Results without spaces (cross)

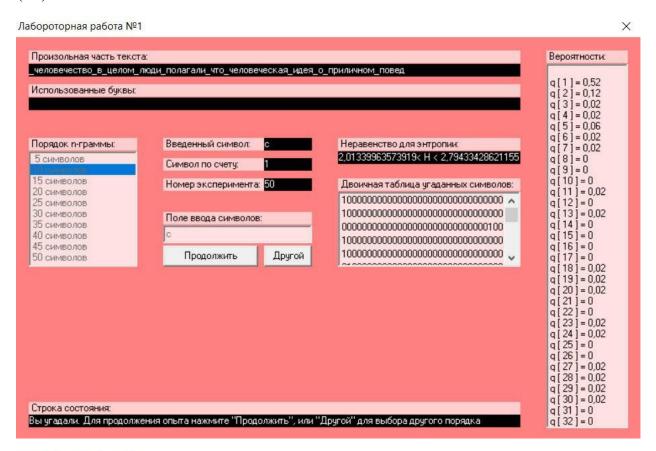
 $a \quad 0.00034 \ 0.00236 \ 0.00527 \ 0.00128 \ 0.00328 \ 0.00190 \ 0.00009 \ 0.00126 \ 0.00476 \ 0.00190 \ 0.00009 \ 0.00127 \ 0.00127 \ 0.00470 \ 0.0079 \ 0.00240 \ 0.00335 \ 0.00316 \ 0.00672 \ 0.00623 \ 0.00069 \ 0.00038 \ 0.00135 \ 0.00059 \ 0.00161 \ 0.00087 \ 0.00044 \ 0.00009 \ 0$ $6 \quad 0.00079 \ 0.00001 \ 0.00007 \ 0.00001 \ 0.00007 \ 0.00001 \$ $\pi = 0.00465\ 0.00009\ 0.00128\ 0.00007\ 0.00029\ 0.00077\ 0.00000\ 0.00007\ 0.00000\ 0.00007\ 0.00000\ 0.00007\ 0.00000\ 0.00007\ 0.00000\ 0.00007\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.000000\ 0.000000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.000000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.000000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.000000\ 0.00000\ 0$ $= 0.00039\ 0.00195\ 0.00284\ 0.00340\ 0.0048\ 0.00171\ 0.00000\ 0.00108\ 0.00221\ 0.00085\ 0.00206\ 0.00289\ 0.00195\ 0.00195\ 0.00195\ 0.00102\ 0.00008\ 0.00102\ 0.00012\ 0.00012\ 0.00012\ 0.0$ $\epsilon = 0.00000, 0.0000$ $3 - 0.00694\ 0.00019\ 0.00138\ 0.00054\ 0.00127\ 0.00058\ 0.00000\ 0.00003\ 0.00001\ 0.00005\ 0.00000\ 0.00003\ 0.00001\ 0.00005\ 0.00000\ 0.00001\ 0.00005\ 0.00000\ 0.00001\ 0.00005\ 0.000000\ 0.00001\ 0.000$ $0.00078\ 0.00120\ 0.00450\ 0.00120\ 0.00450\ 0.00170\ 0.00234\ 0.00076\ 0.00154\ 0.00024\ 0.00000\ 0.00047\ 0.00298\ 0.0024\ 0.000000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.000000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.000000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.000000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.000000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.000000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.000000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.000000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.000000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.000000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.000000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.000000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.000000\ 0.00000\ 0.00000\ 0.00000\ 0.000000\ 0.000000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.000$ $\breve{\textbf{x}} = 0.00012\ 0.00020\ 0.00062\ 0.00020\ 0.00060\ 0.00012\ 0.00000\ 0.00001\ 0.00000\ 0.00000\ 0.00001\ 0.00000\ 0.00001\ 0.00000\ 0.00000\ 0.00001\ 0.00000\ 0.00001\ 0.000000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.000000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.000000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.000000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.000000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.000000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.000000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.000000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.000000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.000000\ 0.00000\ 0.00000\ 0.00000\ 0.000000\ 0.000000\ 0.000000\ 0.00000\ 0.000000\ 0.00000\ 0.00000\ 0.000000\ 0.00000\ 0.00000\ 0.$ $0.00819\ 0.00019\ 0.00046\ 0.00006\ 0.00028\ 0.00110\ 0.00006\ 0.00028\ 0.00110\ 0.00000\ 0.00007\ 0.00014\ 0.00000\ 0.00007\ 0.00001\ 0.00009\ 0.000021\ 0.00009\ 0.00004\ 0.00045\ 0.00241\ 0.00110\ 0.00136\ 0.00227\ 0.00003\ 0.00008\ 0.00010\ 0.00006\ 0.000000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.000000\ 0.000000\ 0.00000\ 0.00000\ 0.00000\ 0.000000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.000$ $0.00367\ 0.00040\ 0.00064\ 0.00019\ 0.00041\ 0.00006\ 0.00007\ 0.000007\ 0.00007\ 0.00007\ 0.00007\ 0.00007\ 0.00007\ 0.00007\ 0.000007\ 0.00007\ 0.00007\ 0.00007\ 0.00007\ 0.00007\ 0.00007\ 0.000007\ 0.00007\ 0.00007\ 0.00007\ 0.00007\ 0.00007\ 0.00007\ 0.000007\ 0.00007\ 0.00007\ 0.00007\ 0.00007\ 0.00007\ 0.00007\ 0.000007\ 0.00007\ 0.00007\ 0.00007\ 0.00007\ 0.00007\ 0.00007\ 0.000007\ 0.00007\ 0.00007\ 0.00007\ 0.00007\ 0.00007\ 0.00007\ 0.000007\ 0.00007\ 0.00007\ 0.00007\ 0.00007\ 0.00007\ 0.00007\ 0.000007\ 0.00007\ 0.00007\ 0.00007\ 0.00007\ 0.00007\ 0.00007\ 0.000007\ 0.00007\ 0.00007\ 0.00007\ 0.00007\ 0.00007\ 0.00007\ 0.000007\ 0.00007\ 0.00007\ 0.00007\ 0.00007\ 0.00007\ 0.00007\ 0.000007\ 0.00007\ 0.00007\ 0.00007\ 0.00007\ 0.00007\ 0.00007\ 0.000007\ 0.00007\ 0.00007\ 0.00007\ 0.00007\ 0.00007\ 0.00007\ 0.000007\ 0.00007\ 0.00007\ 0.00007\ 0.00007\ 0.00007\ 0.00007\ 0.0000$ $0.01418\ 0.00021\ 0.00060\ 0.00021\ 0.00140\ 0.0007\ 0.00000\ 0.00007\ 0.00000\ 0.00007\ 0.00008\ 0.00007\ 0.00008\ 0.00007\ 0.00008\ 0.00007\ 0.00008\ 0.00007\ 0.00008\ 0.00007\ 0.00008\ 0.00007\ 0.00008\ 0.00007\ 0.00008\ 0.00007\ 0.00008\ 0.00007\ 0.00008\ 0.00009\ 0.$ $0.00028\ 0.00508\ 0.00956\ 0.00484\ 0.00524\ 0.00203\ 0.00000\ 0.00232\ 0.00272\ 0.00168\ 0.00396\ 0.00383\ 0.00904\ 0.00761\ 0.00986\ 0.00196\ 0.00409\ 0.00707\ 0.01048\ 0.00890\ 0.00073\ 0.00034\ 0.00107\ 0.00018\ 0.00248\ 0.00099\ 0.00009\ 0.000000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.000000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.000000\ 0.000000\ 0.000000\ 0.000000\ 0.000000\ 0.000000\ 0.000000\ 0.00000\ 0.000000\ 0.000000\ 0.000000\ 0.0000000\ 0.000000\ 0.000000\ 0.00000\ 0.00000\ 0.00000\ 0.000000\ 0.00000\ 0.00000\ 0.00000\ 0.$ $-0.00211\ 0.00001\ 0.00004\ 0.00001\ 0.00004\ 0.00001\ 0.000002\ 0.00020\ 0.000000\ 0.000000\ 0.000000\ 0.000000\ 0.000000\ 0.000000\ 0.000000\ 0.000000\ 0.000000\ 0.000000\ 0.000000\ 0.000000\ 0.000000\ 0.000000\ 0.0$ $p = 0.00961\ 0.00023\ 0.00090\ 0.00023\ 0.00096\ 0.00023\ 0.00005\ 0.00016\ 0.00005\ 0.00016\ 0.00006\ 0.00006\ 0.00006\ 0.00006\ 0.00006\ 0.00006\ 0.00006\ 0.00006\ 0.00006\ 0.00006\ 0.00006\ 0.00006\ 0.00006\ 0.00006\ 0.000006\ 0.000$ $= 0.00720\, 0.00022\, 0.00026\, 0.00009\, 0.00044\, 0.00097\, 0.00004\, 0.00006\, 0.00000\, 0.00006\, 0.00000\, 0.000153\, 0.00033\, 0.00032\, 0.00180\, 0.01454\, 0.00072\, 0.00038\, 0.00158\, 0.000158\, 0.00012\, 0.00005\, 0.00043\, 0.00004\, 0.00000\, 0.00183\, 0.00678\, 0.00014\, 0.00007\, 0.00083\, 0.001$ $0.00033\ 0.00000\ 0$ $x = 0.00049\ 0.00013\ 0.00060\ 0.00006\ 0.00006\ 0.00000\ 0.000000\ 0.000000\ 0.000000\ 0.00000\ 0.00000\ 0.00000\ 0.000000\ 0.000000\ 0.000000\$ ${\bf q} = 0.00063\,0.00001\,0.00008\,0.00001\,0.00008\,0.00001\,0.00000\,0.00$ $= 0.00194\ 0.00000\ 0.00001\ 0.000000$ $0.00105\ 0.000000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.000000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.000000\ 0.000000\ 0.000000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.000$ $0.00045\ 0.00000\ 0$ **5** 0.00000 0. 61 0.00010 0.00059 0.00176 0.00026 0.00039 0.00132 0.00000 0.000012 0.00004 0.00012 0.00004 0.00012 0.00004 0.000012 0.00004 0.000012 0.00004 0.000012 0.00004 0.000012 0.00000 0.00000 0.000001 0.00000 0.000

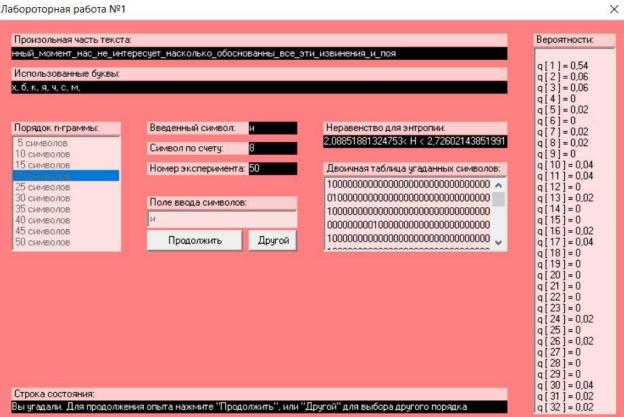
Results without spaces (non-cross)

 $6 \quad 0.00074 \ 0.00001 \ 0.00010 \ 0.00001 \ 0.00001 \ 0.00001 \ 0.00001 \ 0.00002 \ 0.00012 \ 0.00002 \ 0.00013 \ 0.00002 \ 0.00001 \$ $B = 0.00672\ 0.00019\ 0.00042\ 0.00025\ 0.00048\ 0.00046\ 0.00009\ 0.00008\ 0.00008\ 0.00008\ 0.00008\ 0.00008\ 0.00009\ 0.000009\ 0.000009\ 0.000009\ 0.000009\ 0.00009\ 0.00009\ 0.000009\ 0.000009\ 0.00009\ 0.00009\$ $\pi = 0.00455\ 0.00008\ 0.00130\ 0.00007\ 0.00009\ 0.00008\ 0.00010\ 0.00009\ 0.00008\ 0.00009\ 0.00008\ 0.00009\ 0.00008\ 0.00009\ 0.00008\ 0.00009\ 0.000009\ 0.000009\ 0.000009\ 0.000009\ 0.00009\ 0.00009\ 0.000009\ 0.00009\ 0.00009\ 0.00009\ 0$ x 0.00108 0.00008 0.00001 0.00001 0.00004 0.0037 0.00000 0.00001 0.00001 0.00001 0.00000 0.000001 0.00000 0.000001 0.000001 0.00000 0.000000 0.000000 0.000 $3 - 0.00686 \ 0.00018 \ 0.00138 \ 0.00049 \ 0.00138 \ 0.00049 \ 0.00123 \ 0.00054 \ 0.000001 \ 0.00001 \$ $\breve{\textbf{u}} = 0.00014 \ 0.00019 \ 0.00058 \ 0.00020 \ 0.00066 \ 0.00011 \ 0.00007 \ 0.00014 \ 0.00019 \ 0.00058 \ 0.00014 \ 0.00019 \ 0.00$ $0.00740\ 0.00027\ 0.00078\ 0.00036\ 0.00259\ 0.00074\ 0.000074\ 0.00007\ 0.00078\ 0.00036\ 0.00025\ 0.00000\ 0.00003\ 0.000074\ 0.00001\ 0.00068\ 0.00009\ 0.000000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.000000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.000000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.000000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.00000\ 0.000000\ 0.0$ $\ \, \mathbf{M} = 0.00370\ 0.00041\ 0.00061\ 0.00018\ 0.00039\ 0.00040\ 0.00008\ 0.000008\ 0.00008\ 0.00008\ 0.000008\ 0.00008\ 0.00008\ 0.00008\ 0.00008\ 0.00008\ 0.00008\ 0.00$ $\mathbf{H} = 0.01402\ 0.00021\ 0.00061\ 0.00019\ 0.00160\ 0.00019\ 0.00106\ 0.00001\ 0.00001\ 0.00001\ 0.000000\ 0.00008\ 0.000033\ 0.000871\ 0.000000\ 0.00006\ 0.000014\ 0.00009\ 0.00012\ 0.00013\ 0.000467\ 0.00009\ 0.00012\ 0.00001\ 0.00001\ 0.00001\ 0.000010\ 0.00001\ 0.000$ $p = 0.00932\ 0.00022\ 0.00091\ 0.00024\ 0.00036\ 0.00787\ 0.00000\ 0.00052\ 0.000052\ 0.000052\ 0.000052\ 0.000053\ 0.00010\ 0.000053\ 0.00011\ 0.000052\ 0.00015\ 0.00015\ 0.00015\ 0.000015\ 0.00015\ 0.0000015\ 0.000015\ 0.000015\ 0.000015\ 0.000015\ 0.000015\ 0.000015\ 0.000015\ 0.000015\ 0.0$ $y = 0.00018 \ 0.00121 \ 0.00136 \ 0.00121 \ 0.00136 \ 0.00113 \ 0.00196 \ 0.00028 \ 0.00000 \ 0.00144 \ 0.000074 \ 0.00001 \ 0.00019 \ 0.00019 \ 0.00019 \ 0.00013 \ 0.00019 \ 0.00013 \ 0.00013 \ 0.00019 \ 0.00013 \ 0.00013 \ 0.00013 \ 0.00019 \ 0.00013 \ 0.00019 \$ $\mathbf{x} = 0.00051 \ 0.00012 \ 0.00057 \ 0.000$ $\underline{\mathbf{m}} = 0.00097 \ 0.00007 \ 0.0$

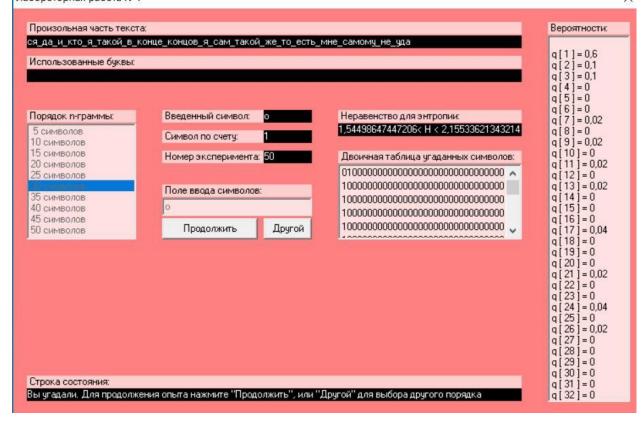
6.1 0.00010 0.00055 0.00180 0.00028 0.00041 0.00129 0.00000 0.00010 0.00003 0.00011 0.00003 0.00011 0.00003 0.00011 0.00003 0.00011 0.00000 0

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





Лабороторная работа №1



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

Results with spaces:

• R letter: 0.1372150079813922

• R cross bigrams: 0.21376879449767883

• R non-cross bigrams: 0.2143173515988962

Results without spaces:

• R letter: 0.1163321751646319

• R cross bigrams: 0.17768564958330402

• R non-cross bigrams: 0.177766347908082

Results for CoolPinkProgramm:

• 0.44113314275769 < R(10) < 0.597320072852162

• 0.454795712296018 < R(20) < 0.5822962373504941

• 0.568932757313572 < R(30) < 0.691002705105588

Висновки:

В результаті виконання даної лабораторної роботи ми отримали багато навичок, а саме:

- о Підрахування частоти монограм і біграм у тексті
- о Знаходження ентропії для монограм і біграм
- о Визначення надлишковості мови
- Робота з CoolPinkProgram

При виконанні 1 завдання (а саме той пункт, де вираховували частоту монограм), ми практично переконалися у тому, що літерами з найбільшими частотами ϵ 'o', 'a' та 'e'. В той час як літерами з найменшою частотою ϵ 'ъ' та 'ë'.

В 3 пункті лабораторної роботи ми вирахували надлишковість російської мови, побачили її залежність від ентропії та кількості символів в алфавіті.