НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ УКРАЇНИ

КИЇВСЬКИЙ ПОЛІТЕХНІЧНИЙ ІНСТИТУТ ІМЕНІ ІГОРЯ СІКОРСЬКОГО

ФІЗИКО-ТЕХНІЧНИЙ ІНСТИТУТ

КАФЕДРА ІНФОРМАЦІЙНОЇ БЕЗПЕКИ

Комп’ютерний практикум  
з дисципліни

«КРИПТОГРАФІЯ»

Лабораторна №1

Виконав: ФБ-06 Березовський М.Ю.

Перевірив:

Київ – 2022

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 комп'ютерного практикуму

1. Уважно прочитати методичні вказівки до виконання комп’ютерного

практикуму.

2. Написати програми для підрахунку частот букв і частот біграм в тексті, а також підрахунку H1 та H2 за безпосереднім означенням. Підрахувати частоти букв та біграм, а також значення H1 та H2 на довільно обраному тексті російською мовою достатньої довжини (щонайменше 1Мб) , де імовірності замінити відповідними частотами. Також одержати значення H 1 та H2 на тому ж тексті, в якому вилучено всі пробіли.

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

4.Використовуючи отримані значення ентропії, оцінити

надлишковість російської мови в різних моделях джерела

Хід роботи

Під час виконання лабораторної роботи використовував написаний мною код на Python, збережений під назвою **main.py** .   
 Код має чітку структуру та поділений на функції:

freq\_letter

-частота символів

freq\_bigram

-Рахує частоту біграм у тексті

bigram\_freq\_cross

-частота у крос біграмах

entropy

-ентропія

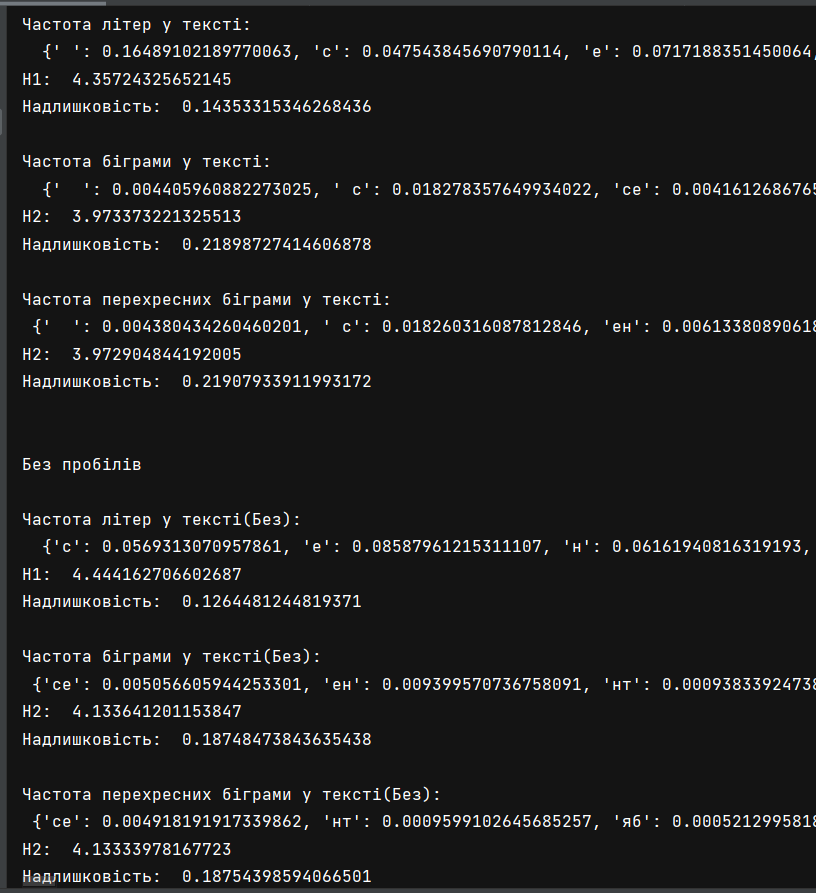
Також стикався із незначними проблемами, не зразу зрозумів принципи роботи програми CoolPinkProgram.  
Також доволі складно писати код та виконувати завдання коли по 5-6 разів на день вимикається світло і втрачається прогрес написання та мотивація.   
Складна формульно-математична частина лабораторної, але кілька відео на ютубі та практика роблять свою справу.

Значення отримані в результаті виконання роботи

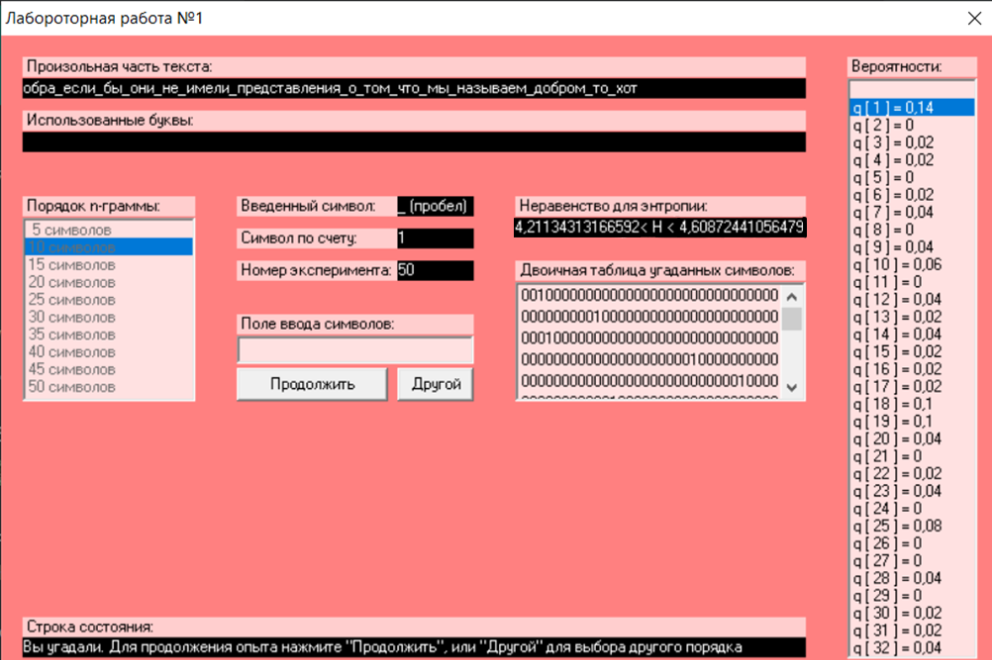
|  |  |
| --- | --- |
| **Значення з пробілами зеленого кольору** | **Frequency of letters in samptext:** |
| ' ' | 0.16489102189770063 |
| 'с' | 0.047543845690790114 |
| 'е' | 0.0717188351450064 |
| 'н' | 0.0514589209824317 |
| 'т' | 0.054390723913634156 |
| 'я' | 0.01671863248581761 |
| 'б' | 0.013343981697623483 |
| 'р' | 0.034905418206000814 |
| 'о' | 0.09429056523899523 |
| 'к' | 0.03091978749307581 |
| 'г' | 0.012907138557336788 |
| 'а' | 0.07386852372538628 |
| 'м' | 0.024782967020594678 |
| 'в' | 0.03621894998521338 |
| 'ц' | 0.002997854814957149 |
| 'п' | 0.021803126355752444 |
| 'л' | 0.040291649090360475 |
| 'ь' | 0.01738665721924572 |
| 'й' | 0.008759380493377548 |
| 'ч' | 0.012829077240172156 |
| 'и' | 0.051750149742622834 |
| 'ы' | 0.014774605452583004 |
| 'у' | 0.024803983529062077 |
| 'э' | 0.0026435765293638153 |
| 'д' | 0.02746557420854081 |
| 'ж' | 0.008947027890407916 |
| 'х' | 0.005683464361255767 |
| 'з' | 0.015975548793577356 |
| 'ш' | 0.006280933673400456 |
| 'ю' | 0.0050244467028851166 |
| 'щ' | 0.003045892548596923 |
| 'ъ' | 0.00019665447208782499 |
| 'ф' | 0.001324040033446272 |
| 'ё' | 5.7044808697231676e-05 |

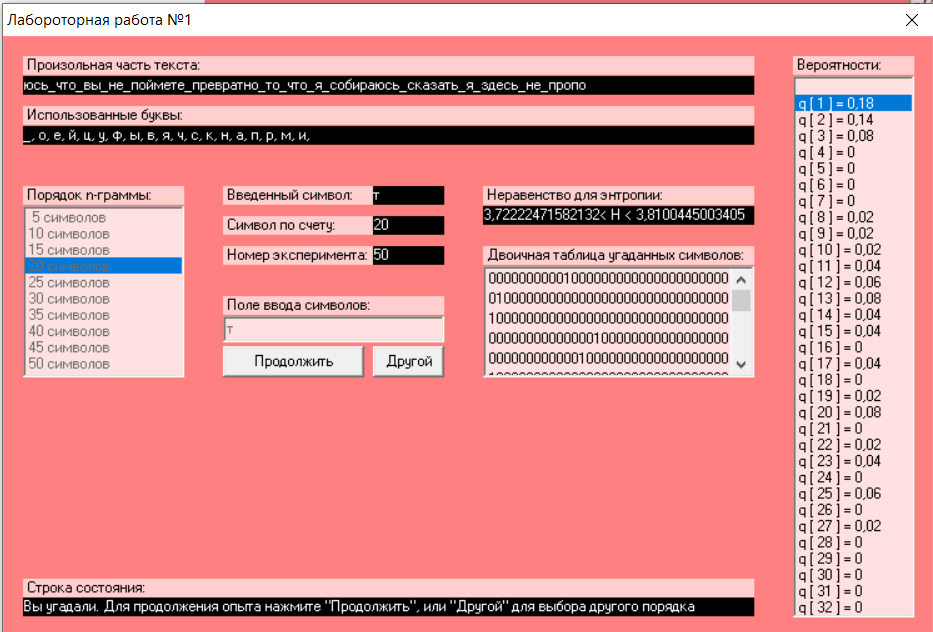
|  |
| --- |
| **Frequency of bigrams in samptext:** |
| ' ': 0.004405960882273025 |
| ' с': 0.018278357649934022 |
| 'се': 0.0041612686765454265 |
| 'ен': 0.006078274484607659 |
| 'нт': 0.0006500105833131925 |
| 'тя': 0.000586961057910989 |
| 'яб': 6.455070457844636e-05 |
| 'бр': 0.000956251135266752 |
| 'ря': 0.001127385561358447 |
| 'я ': 0.009777179974870261 |
| ' о': 0.010975120957512127 |
| 'ок': 0.0024694397449196346 |
| 'ко': 0.008615267292458225 |
| 'ог': 0.003874543453883025 |
| 'га': 0.0012354704620479386 |
| 'ам': 0.003562298185224494 |
| 'ма': 0.0029348052895549455 |
| 'а ': 0.01765837065014569 |
| 'ев': 0.0012639928663965546 |
| 'ве': 0.00469418728411167 |
| 'ер': 0.0065361341333617555 |
| 'рн': 0.0010673383943087295 |
| 'но': 0.010664376868029837 |
| 'ом': 0.004484022199437658 |
| 'м ': 0.0063740067823275185 |
| ' к': 0.007741581011884835 |
| 'он': 0.0048908417561994945 |
| 'нц': 0.000220673338907712 |
| 'це': 0.000687540062719266 |
| 'е ': 0.018754231448803036 |
| 'сп': 0.001739866665265566 |
| 'па': 0.0015552216265876846 |
| 'ал': 0.0074998911645097225 |
| 'ль': 0.003900063499879155 |
| 'ьн': 0.0013930942755534471 |
| 'го': 0.006040745005201586 |
| 'яц': 6.755306293093224e-05 |
| 'йш': 7.505895881214695e-05 |
| 'юл': 6.004716704971755e-06 |
|  |
|  |
|  |
|  |
|  |
|  |
|  |
|  |
|  |
|  |
|  |
| |  |  | | --- | --- | | Значення без пробілів |  | |
| |  | | --- | | **Frequency of letters in samptext:** | | с': 0.0569313070957861 | | 'е': 0.08587961215311107 | | 'н': 0.06161940816319193 | | 'т': 0.06513009120945099 | | 'я': 0.020019701529025602 | | 'б': 0.015978730977059222 | | 'р': 0.04179744095832839 | | 'о': 0.112908096681299 | | 'к': 0.037024853406962405 | | 'г': 0.015455633810412329 | | 'а': 0.08845375353674803 | | 'м': 0.029676326887194366 | | 'в': 0.04337032762779929 | | 'ц': 0.003589776775923869 | | 'п': 0.026108121128451813 | | 'л': 0.04824717509554163 | | 'ь': 0.020819626749499373 | | 'й': 0.010488907104414509 | | 'ч': 0.01536215940262663 | | 'и': 0.0619681396076232 | | 'ы': 0.017691829258208672 | | 'у': 0.0297014930739059 | | 'э': 0.0031655467713580035 | | 'д': 0.03288861086244522 | | 'ж': 0.010713605200053208 | | 'х': 0.006805655920704941 | | 'з': 0.01912989707029635 | | 'ш': 0.0075210946572185616 | | 'ю': 0.006016516208821827 | | 'щ': 0.003647299488407376 | | 'ъ': 0.00023548360422935745 | | 'ф': 0.0015854697628266661 | | 'ё': 6.830822107416475e-05 | |

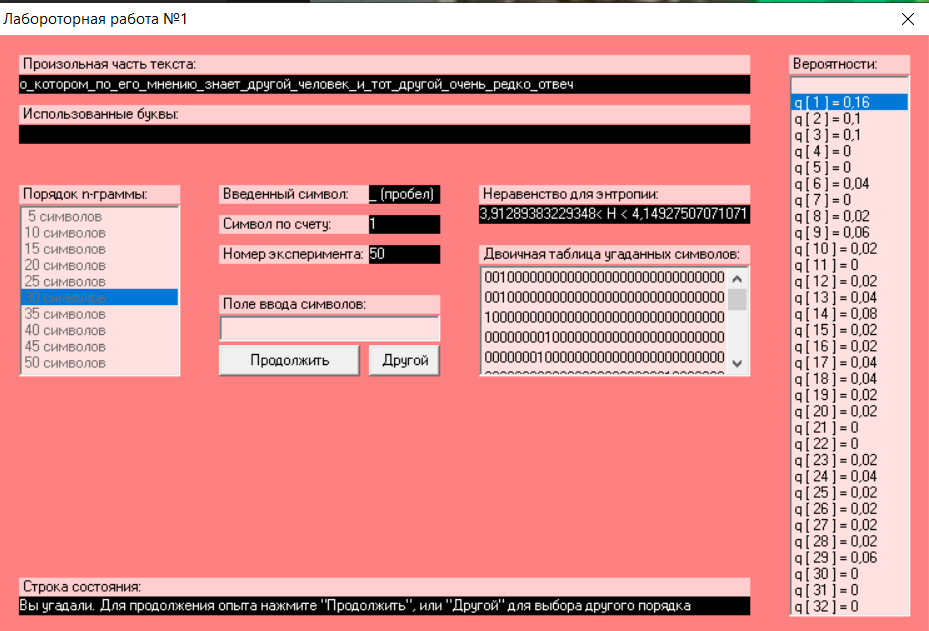
Повна інформація після виконання коду



Cool Pink Program

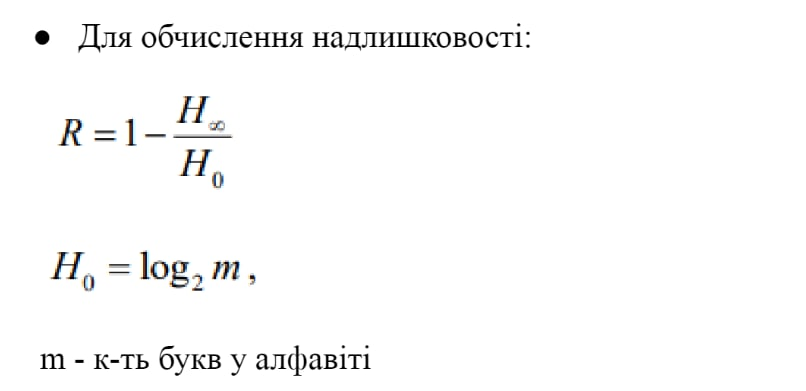






|  |  |  |
| --- | --- | --- |
| Cool Pink Program | | |
|  | H0 | R |
| H10 | 4,410032 | 0,117994 |
| H20 | 3,76613 | 0,246774 |
| H30 | 4,03108 | 0,193784 |

Отримуємо значення після розрахунків за формулою даною у методичці, а саме:



Висновки:   
Під час лабораторної роботи було здобуто практичні навички використання ентропії та надлишковості символьного джерела, вивчив та порівняв різні моделі відкритого тексту для наближення значень ентропії . Також визначив, що більше всього використовують у російській мові букви : о, а, е, и, н.