**МІНІСТЕРСТВО ОСВІТИ І НАУКИ, МОЛОДІ ТА СПОРТУ 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 гр. ФБ-14

Цибулено-Сігов І. М.

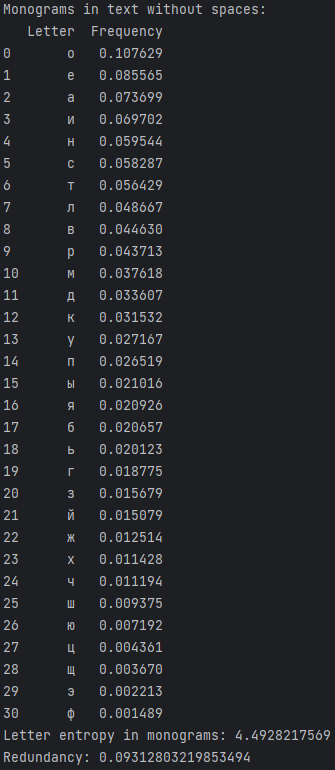
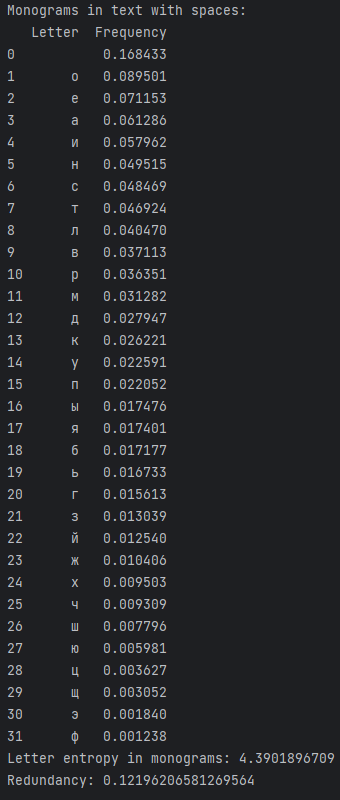
Татаренко А. О.

Перевірила

Селюх П. В.

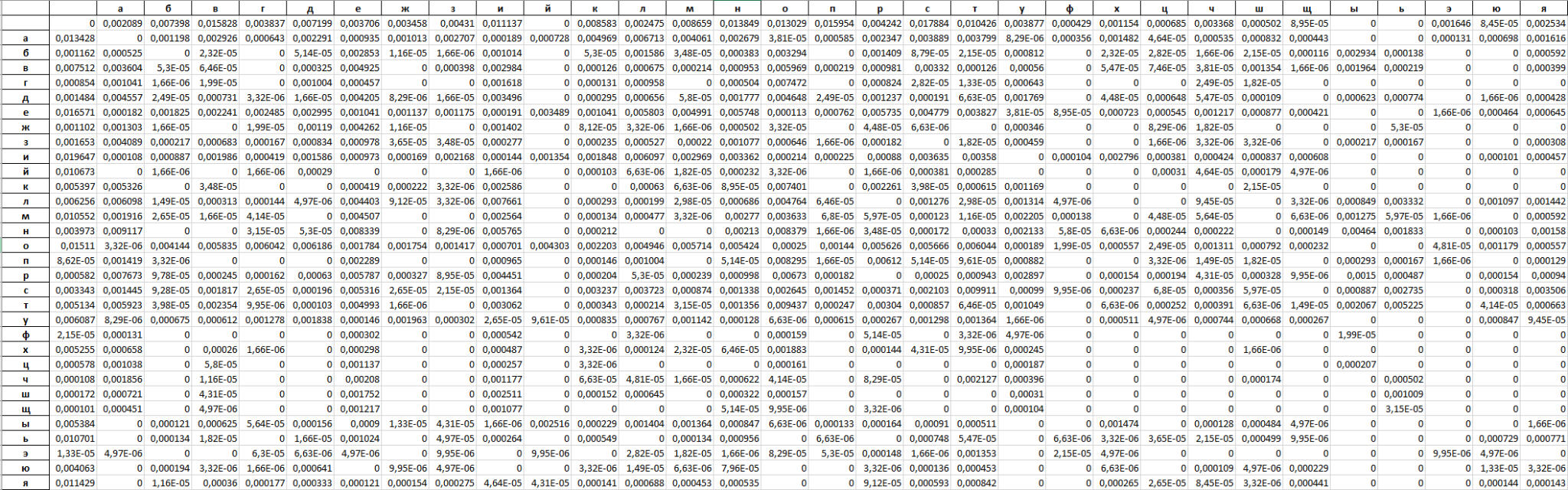
Київ 2023

1. 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 H1 та H2 на тому ж тексті, в якому вилучено всі пробіли.



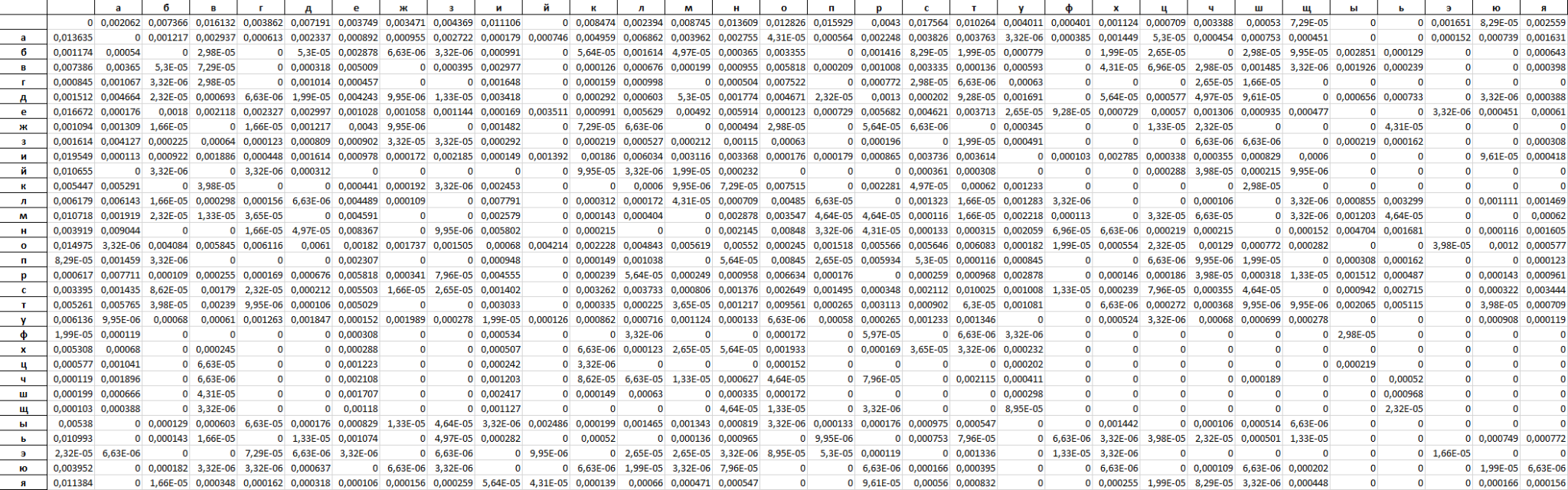
*Примітка: всі таблиці є на гітхабі*

Text with spaces and with intersection



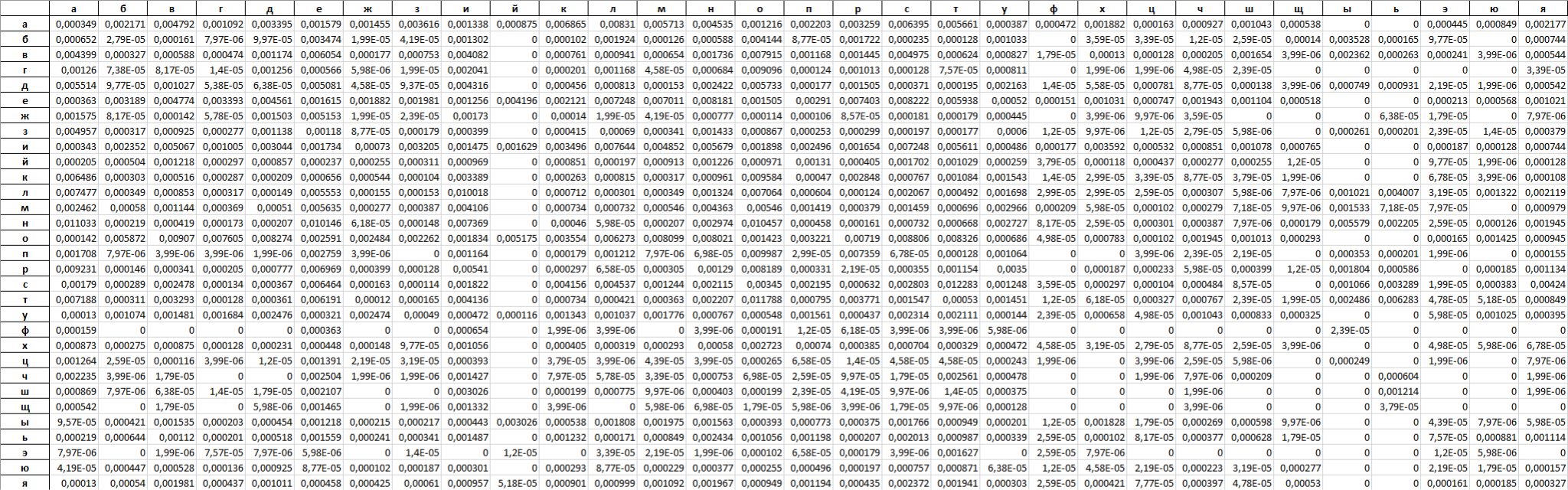


Text with spaces and without intersection



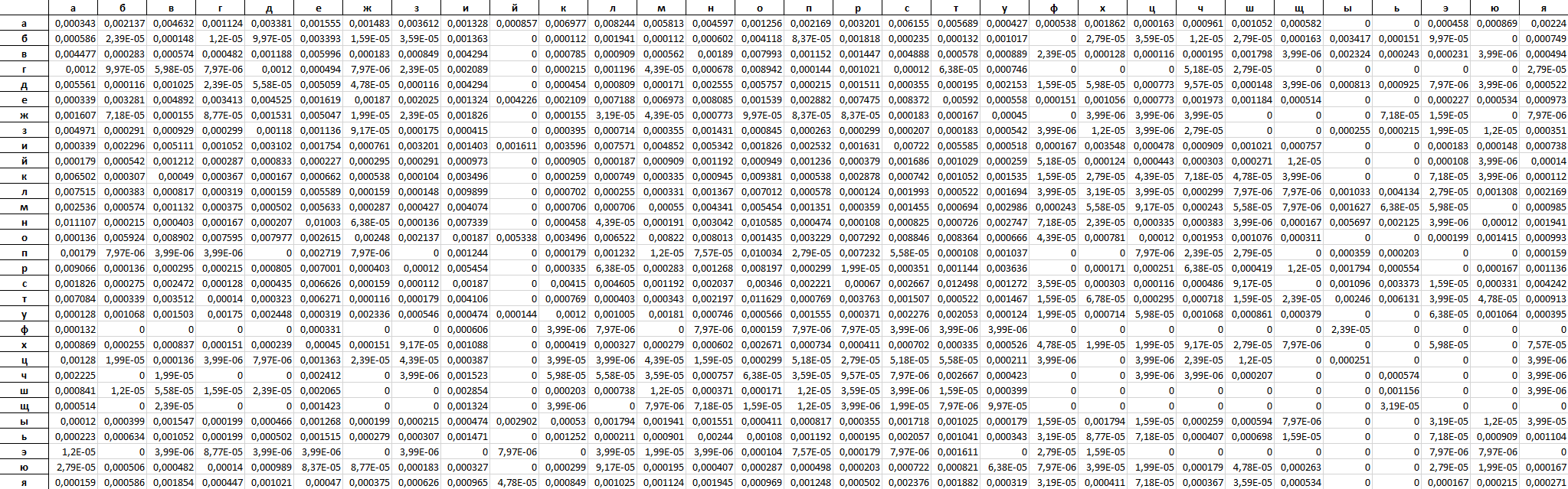


Text without spaces and with intersection





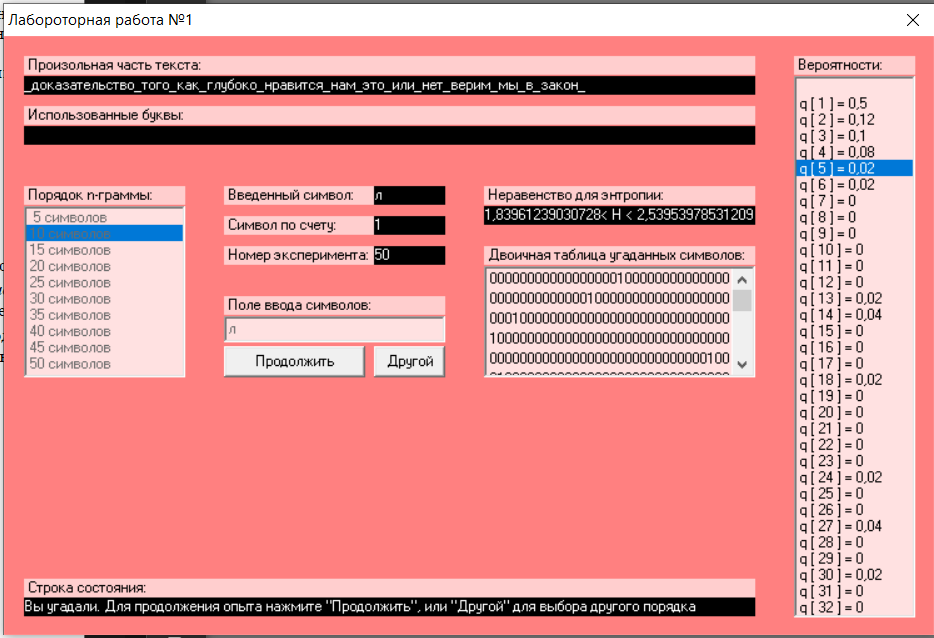
Text without spaces and without intersection





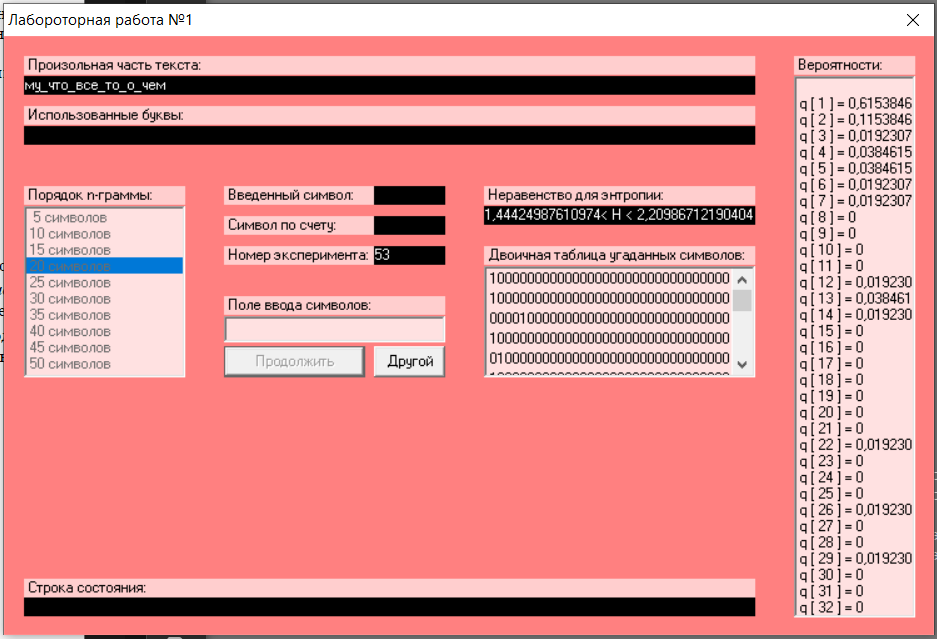
2. За допомогою програми CoolPinkProgram оцінити значення H10, H20, H30.

H10:



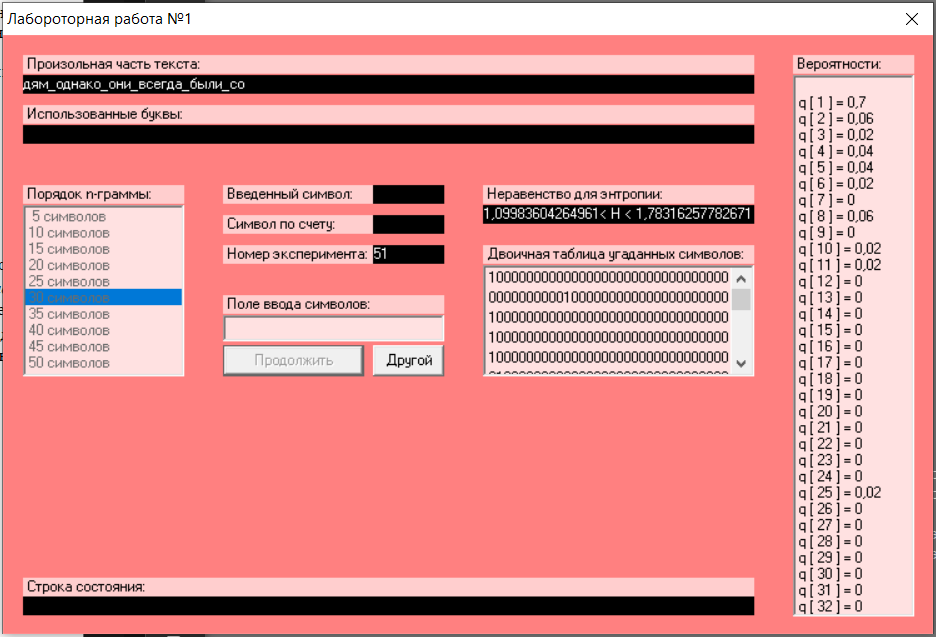
1,8396123903 > H10 > 2,5395397853

H20:



1.4442498761>H20>2.2098671219

H30:



1.0998360426>H30>1.7831625778

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

,

H0 = log2(32) = 5

H10:

1)

2)

H20:

1)

2)

H30:

1)

2)

**Висновки:**

Ми засвоїли теоретичні поняття ентропії на символ джерела та його надлишковості.  
Навчилися на практиці визначати частоти монограм та біграм відкритого тексту, а також розраховувати відповідні значення ентропій (Н1, Н2). Попрацювавши з CoolPinkProgram, ми отримали значення ентропій для Н10, Н20, Н30. Використовуючи отримані значення ентропії, оцінили надлишковість російської мови в різних моделях джерела.