**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-33

Ольшевський Б.

Степура Н.

Перевірила

Селюх П. В.

Київ 2025

**Мета роботи**

Засвоєння понять ентропії на символ джерела та його надлишковості, вивчення та порівняння

різних моделей джерела відкритого тексту для наближеного визначення ентропії, набуття

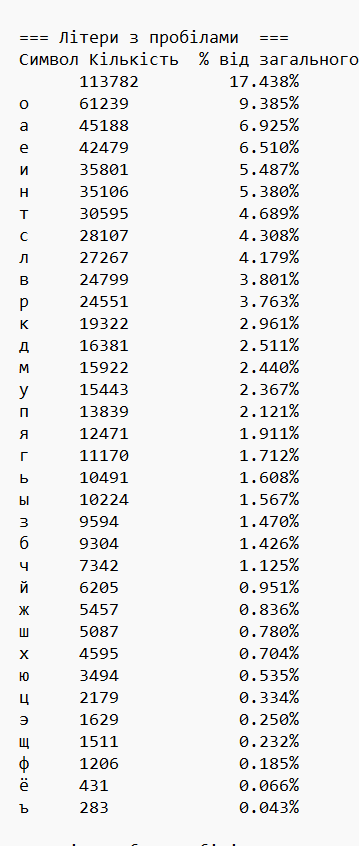
практичних навичок щодо оцінки ентропії на символ джерела.

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.

***Обраний текст - Лев Товстий «Війна і Мир» Том 1 (1.23Мб)***

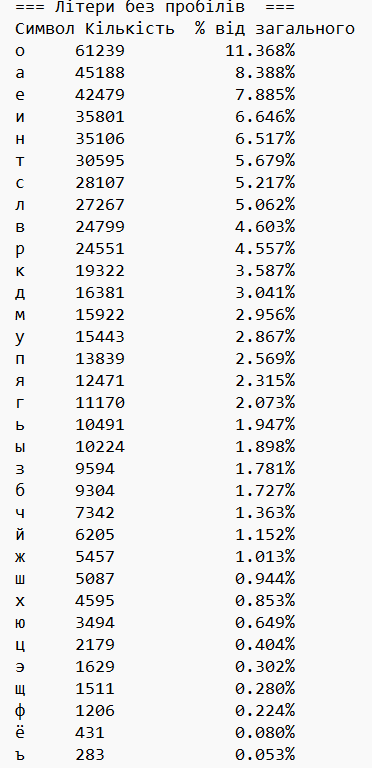
**Букви:**

***З пробілами***



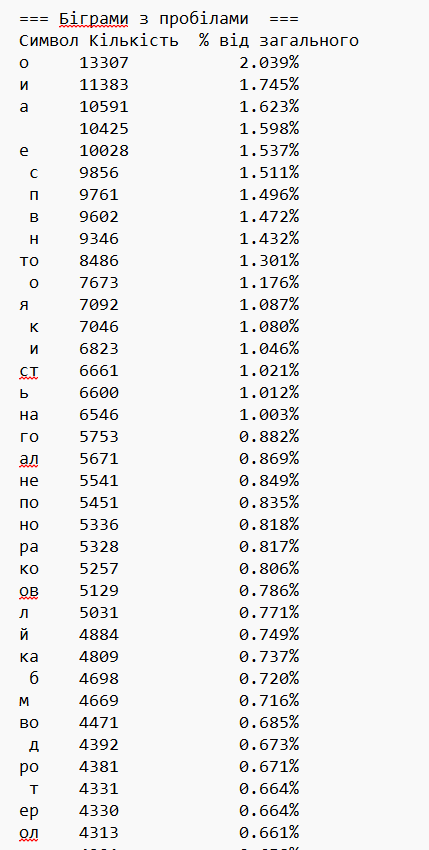
***Без пробілів***

Буква, частота



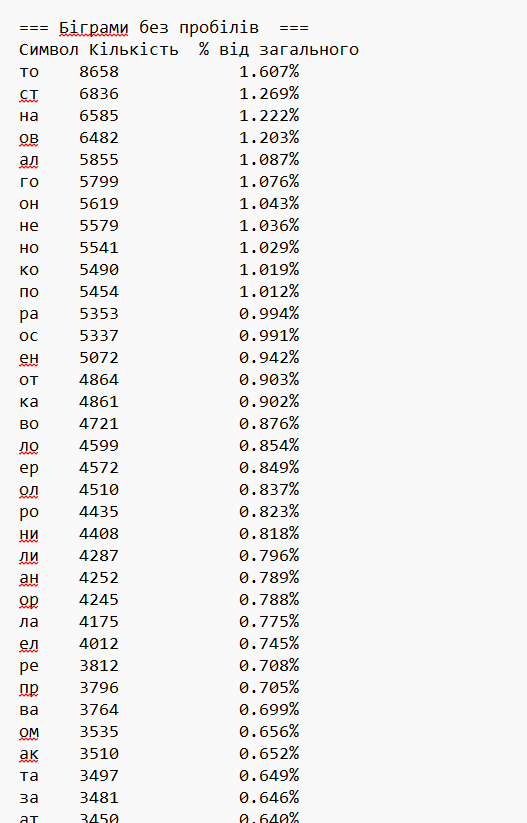
**Біграми, що перетинаються:**

***З пробілами***

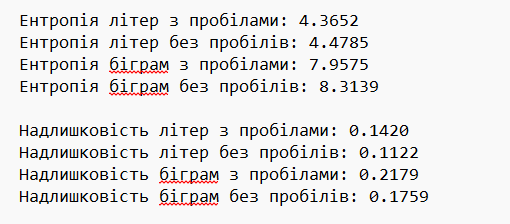
****

Повний файл  *analysis\_results.txt*

***Без пробілів***

****

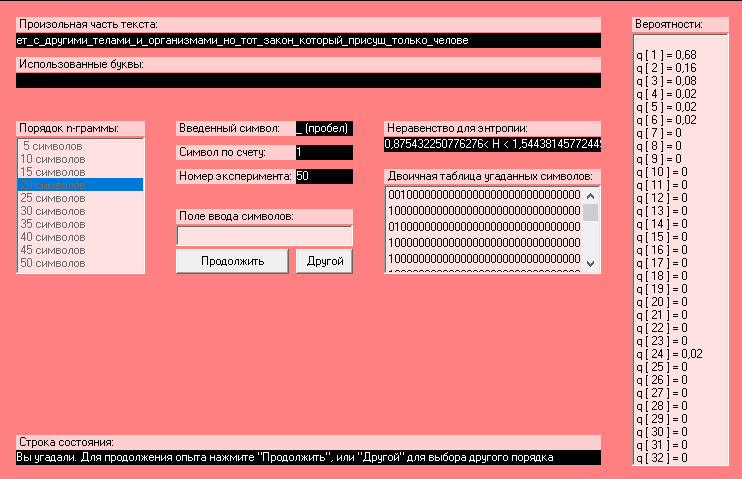
Повний файл  *analysis\_results.txt*

**

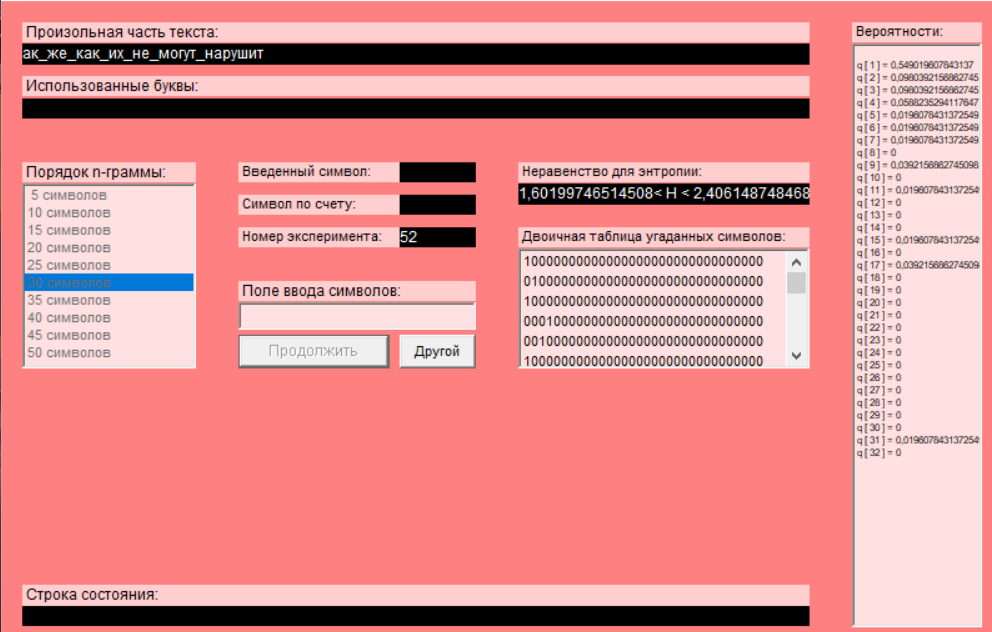
**2. За допомогою програми CoolPinkProgram оцінити значення H10, H20, H30.**



**1,63473262855352 < H(10) < 2,35682542858739**

****

**0,875432250776276 < H(20) < 1,54438145772445**

****

**1,60199746514508 < H(30) < 2,406148748468**

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

-ентропія відкритих текстів. Оскільки відкритий текст в нас рівно вірогідний звідси:

**Висновки:** під час роботи над даним комп’ютерним практикумом ми мали змогу дослідити поняття ентропії, біграм та надлишковості російської мови за допомогою власноруч написаного скрипту та програми CoolPinkProgram, що дало змогу оцінити різницю вихідних даних на практиці у порівнянні з отриманими внаслідок експериментального «вгадування» наступних літер(людського фактору у тому числі)