**Міністерство освіти і науки 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:

Кравченко Владислав Жмур Назар

Група: ФБ-04

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 роботи

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

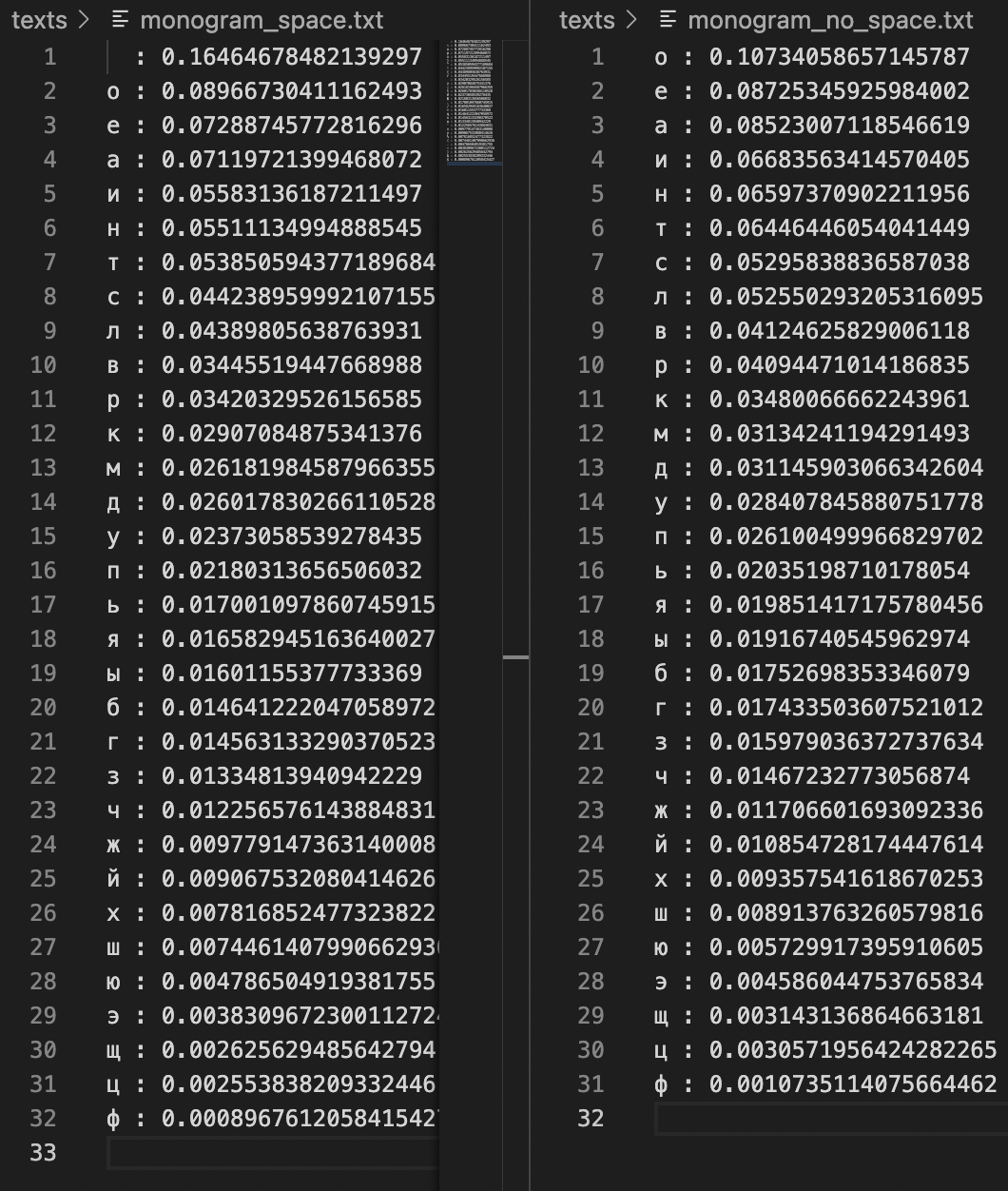
Використовуючи отримані значення ентропії, оцінити надлишковість російської мови в різних моделях джерела.

# Хід роботи

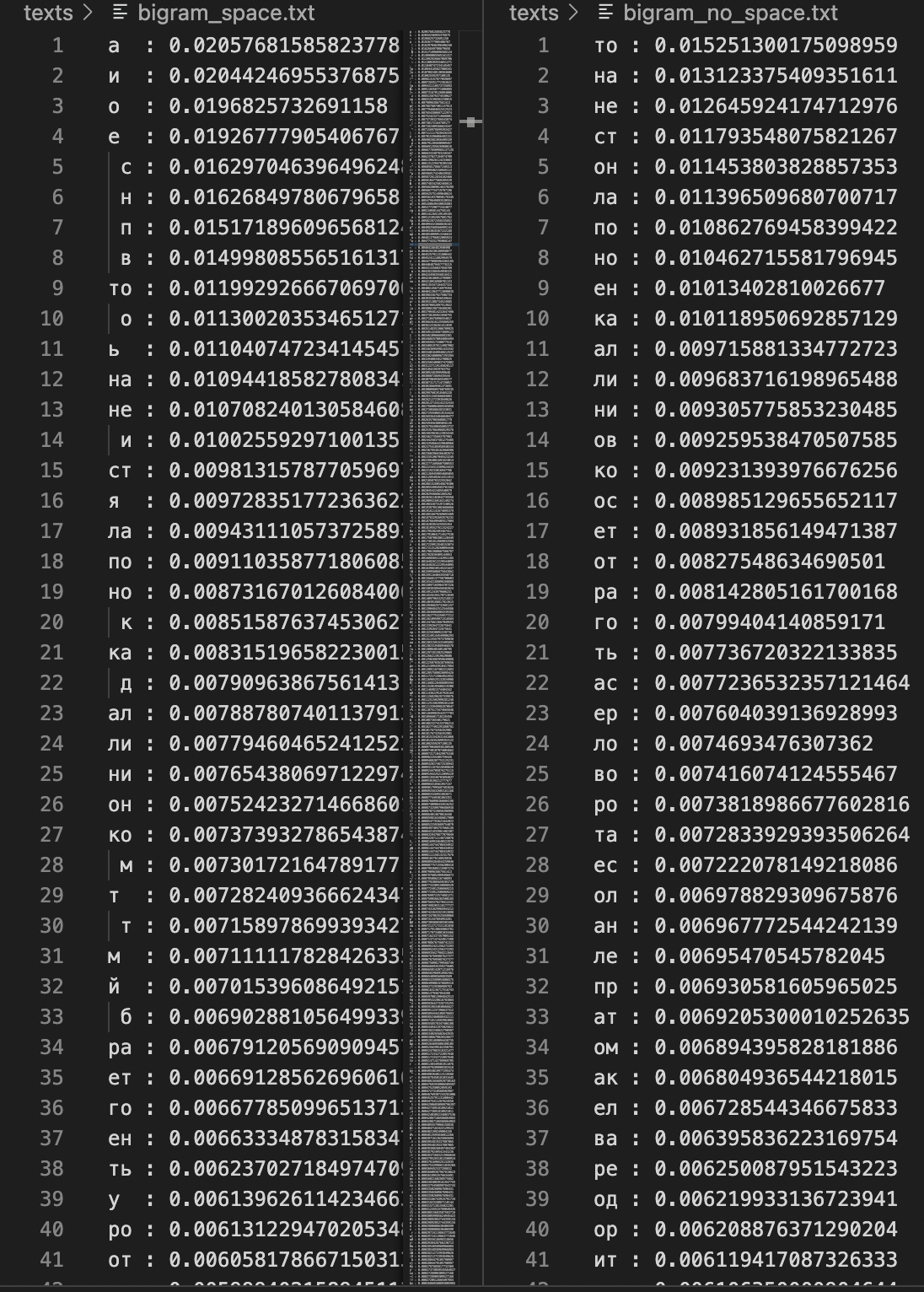
Під час виконання роботи для читання файлу використовували функцію open()

Всі результати виконання роботи були виписані в окремі файли. Далі наведені таблиці з даними підписані та знаходяться нижче.

Частота монограм з та без пробілів



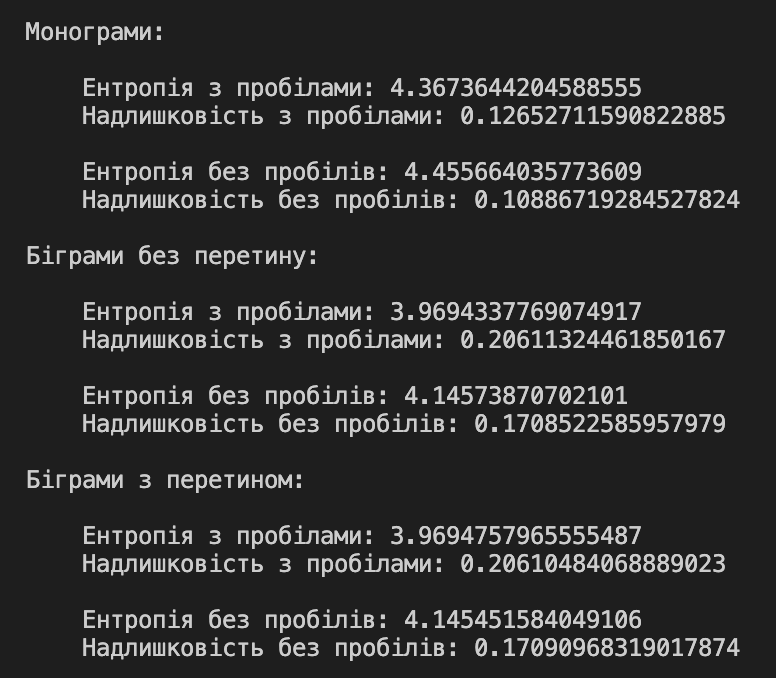
Частота неперехресних біграм з та без пробілів



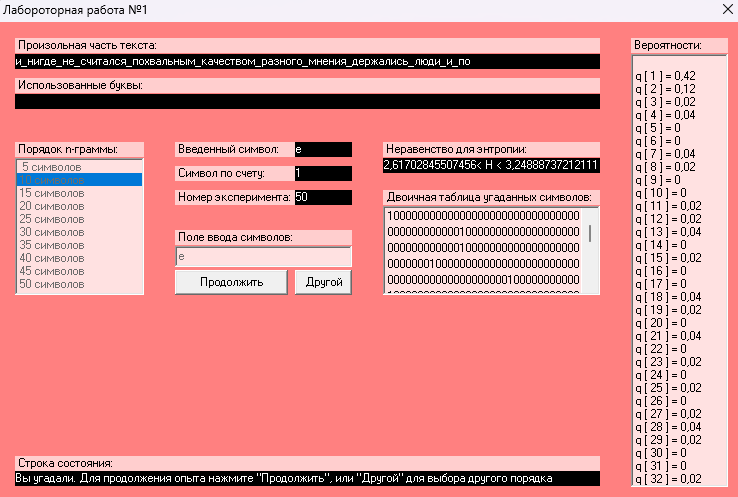
Частота перехресних біграм з та без пробілів



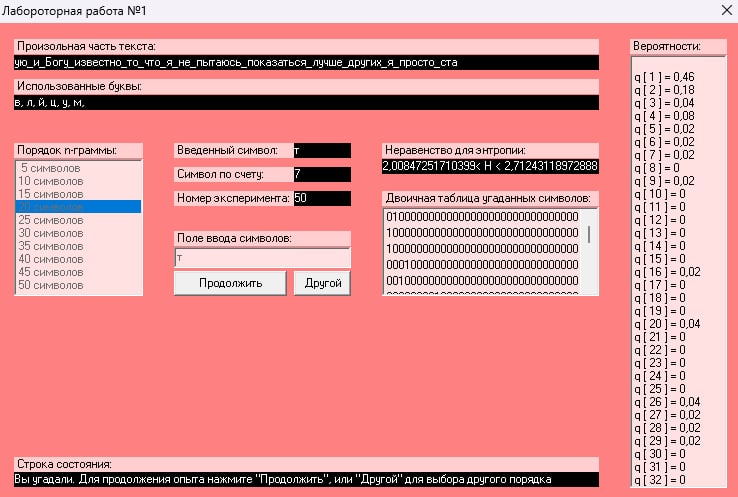
Ентропія та надлишковість



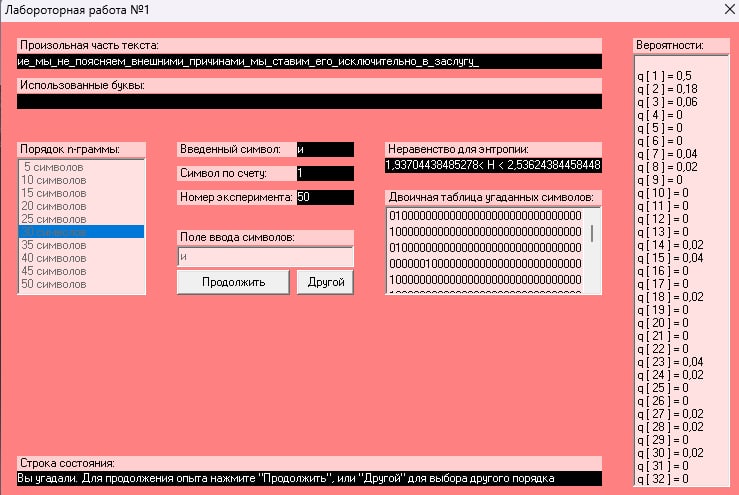
coolpinkprogram



R = 0.41341



R = 0.52792



R = 0.55268

## Висновок

Під час роботи ми навчилися рахувати ентропію, та надлишковість російської мови на прикладі вибраного тексту. З отриамних даних можемо зробити висновок що найчастіше у нашому тексті зустрічаються букви «о», «е» та «а»