



Міністерство освіти і науки України
Національний технічний університет України
«Київський політехнічний інститут»

КРИПТОГРАФІЯ

КОМП'ЮТЕРНИЙ ПРАКТИКУМ №1

Експериментальна оцінка ентропії на символ джерела відкритого тексту

Виконав студент групи: ФБ-24

ПІБ: Мартинюк І. О.

Київ 2024

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.

Методичні вказівки

Звичайні текстові файли містять багато символів окрім власне літер; для обчислення значень ентропій вони повинні пройти попередню фільтрацію: всі символи, окрім текстових, повинні вилучатись або замінюватись на пробіли; прописні літери - замінюватись на відповідні стрічні; послідовність пробілів (або інших розділових знаків, наприклад, символів кінця рядку) повинна трактуватись як один пробіл або вилучатись, якщо пробіл не входить до алфавіту.

При підрахунку частот біграм треба розглядати як пари букв, що перетинаються, такі пари букв, що не перетинаються (тобто рухатися вздовж тексту з кроком 2). Одержані результати не повинні суттєво відрізнятись, однак в першому випадку використовується більше статистики, а тому чисельні дані більш точні. Таблицю частот символів потрібно подавати відсортованою за спаданням частот. Таблицю частот біграм зручно подавати у вигляді квадратної матриці, індексованої першою та другою літерами біграм.

Програма CoolPinkProgram використовує текст, що лежить у допоміжному файлі text. Цей текст написаний російською мовою без знаків пунктуації та великих літер; буква «ё» замінена буквою «е», а «ъ» – буквою «ь». Пробіл також вважається буквою. Таким чином, кількість букв алфавіту $m=32$. При підрахунку $H(10)$, $H(20)$, $H(30)$ виконати не менш ніж 50 експериментів.

Порядок виконання роботи

0. Уважно прочитати методичні вказівки до виконання комп'ютерного практикуму.
1. Написати програми для підрахунку частот букв і частот біграм в тексті, а також підрахунку H_1 та H_2 за безпосереднім означенням. Підрахувати частоти букв та біграм, а також значення H_1 та H_2 на довільно обраному тексті російською мовою достатньої довжини (щонайменше 1Мб), де імовірності замінити відповідними частотами. Також одержати значення H_1 та H_2 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

Частота

Монограми

| 1 | Letter | Count | Frequency |
|----|--------|-------|-----------|
| 2 | | 95597 | 0.1647 |
| 3 | о | 62943 | 0.1084 |
| 4 | е | 50226 | 0.0865 |
| 5 | и | 38751 | 0.0667 |
| 6 | т | 37108 | 0.0639 |
| 7 | н | 32739 | 0.0564 |
| 8 | а | 28998 | 0.0499 |
| 9 | с | 26233 | 0.0452 |
| 10 | в | 20096 | 0.0346 |
| 11 | м | 17450 | 0.0301 |
| 12 | д | 17095 | 0.0294 |
| 13 | л | 16734 | 0.0288 |
| 14 | р | 16495 | 0.0284 |
| 15 | к | 13353 | 0.023 |
| 16 | у | 12471 | 0.0215 |
| 17 | п | 10817 | 0.0186 |
| 18 | ч | 10680 | 0.0184 |
| 19 | ь | 8312 | 0.0143 |
| 20 | я | 8311 | 0.0143 |
| 21 | б | 7969 | 0.0137 |
| 22 | ы | 7888 | 0.0136 |
| 23 | з | 7166 | 0.0123 |
| 24 | г | 6988 | 0.012 |
| 25 | ж | 6487 | 0.0112 |
| 26 | щ | 4436 | 0.0076 |
| 27 | х | 4009 | 0.0069 |
| 28 | й | 3402 | 0.0059 |
| 29 | ю | 2523 | 0.0043 |
| 30 | э | 2280 | 0.0039 |
| 31 | ш | 1748 | 0.003 |
| 32 | ц | 723 | 0.0012 |
| 33 | ф | 448 | 0.0008 |
| 34 | ъ | 107 | 0.0002 |

Рисунок 1: Монограми з пробілами

| 1 | Letter | Count | Frequency |
|----|--------|-------|-----------|
| 2 | о | 62943 | 0.1298 |
| 3 | е | 50226 | 0.1036 |
| 4 | и | 38751 | 0.0799 |
| 5 | т | 37108 | 0.0765 |
| 6 | н | 32739 | 0.0675 |
| 7 | а | 28998 | 0.0598 |
| 8 | с | 26233 | 0.0541 |
| 9 | в | 20096 | 0.0414 |
| 10 | м | 17450 | 0.036 |
| 11 | д | 17095 | 0.0352 |
| 12 | л | 16734 | 0.0345 |
| 13 | р | 16495 | 0.034 |
| 14 | к | 13353 | 0.0275 |
| 15 | у | 12471 | 0.0257 |
| 16 | п | 10817 | 0.0223 |
| 17 | ч | 10680 | 0.022 |
| 18 | ь | 8312 | 0.0171 |
| 19 | я | 8311 | 0.0171 |
| 20 | б | 7969 | 0.0164 |
| 21 | ы | 7888 | 0.0163 |
| 22 | з | 7166 | 0.0148 |
| 23 | г | 6988 | 0.0144 |
| 24 | ж | 6487 | 0.0134 |
| 25 | щ | 4436 | 0.0091 |
| 26 | х | 4009 | 0.0083 |
| 27 | й | 3402 | 0.007 |
| 28 | ю | 2523 | 0.0052 |
| 29 | э | 2280 | 0.0047 |
| 30 | ш | 1748 | 0.0036 |
| 31 | ц | 723 | 0.0015 |
| 32 | ф | 448 | 0.0009 |
| 33 | ъ | 107 | 0.0002 |

Рисунок 2: Монограми без пробілів

Біграми

| | а | б | в | г | д | е | ж | з | и | й | к | л | м | н | о | п | р | с | т | у | ф | х | ц | ч | ш | щ | ъ | ы | ь | э | ю | я | | |
|---|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|---|
| а | 0 | 0.0032 | 0.006 | 0.0142 | 0.0014 | 0.0082 | 0.0056 | 0.0028 | 0.0021 | 0.0175 | 0 | 0.0073 | 0.0019 | 0.0048 | 0.0158 | 0.014 | 0.0133 | 0.003 | 0.0141 | 0.0123 | 0.0026 | 0.0005 | 0.0004 | 0.0004 | 0.0089 | 0.0001 | 0 | 0 | 0 | 0 | 0.0036 | 0 | 0.0006 | |
| б | 0.0102 | 0 | 0.0001 | 0.0019 | 0.0006 | 0.0009 | 0.0033 | 0.0009 | 0.0041 | 0.0002 | 0.0002 | 0.0072 | 0.0027 | 0.0021 | 0.0024 | 0 | 0.0009 | 0.0003 | 0.0017 | 0.0038 | 0.0004 | 0 | 0.0006 | 0 | 0.0022 | 0.0001 | 0.0002 | 0 | 0 | 0 | 0 | 0.0014 | 0.0016 | |
| в | 0.0003 | 0.0002 | 0 | 0 | 0 | 0 | 0.0012 | 0 | 0 | 0.0001 | 0 | 0 | 0.0006 | 0 | 0.0007 | 0.0028 | 0 | 0.0012 | 0.0004 | 0 | 0.0011 | 0 | 0.0005 | 0 | 0 | 0 | 0.0006 | 0.0001 | 0.0037 | 0 | 0 | 0 | 0.0001 | |
| г | 0.0003 | 0.0006 | 0 | 0 | 0 | 0.0008 | 0.0001 | 0.0007 | 0 | 0.0008 | 0 | 0.0001 | 0.0007 | 0 | 0.0001 | 0.0081 | 0 | 0.0002 | 0 | 0 | 0.0003 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | |
| д | 0.0004 | 0.0029 | 0 | 0.0016 | 0 | 0 | 0.0058 | 0 | 0 | 0.0032 | 0 | 0.0001 | 0.0011 | 0.0003 | 0.0035 | 0.0032 | 0.0001 | 0.0024 | 0.0002 | 0.0001 | 0.0019 | 0 | 0 | 0.0001 | 0 | 0.0002 | 0 | 0 | 0.0004 | 0.0011 | 0 | 0 | 0.0008 | |
| е | 0.0248 | 0 | 0.0008 | 0.0009 | 0.0025 | 0.0056 | 0.003 | 0.0012 | 0.0005 | 0.0002 | 0.0019 | 0.0015 | 0.006 | 0.0046 | 0.0084 | 0.0006 | 0.0006 | 0.0046 | 0.0076 | 0.0073 | 0 | 0 | 0.0006 | 0.0001 | 0.0013 | 0.0002 | 0.0013 | 0 | 0 | 0 | 0 | 0.0004 | 0.0002 | |
| ж | 0 | 0.0005 | 0 | 0 | 0 | 0.0019 | 0.0055 | 0 | 0 | 0.0007 | 0 | 0 | 0 | 0 | 0 | 0.0022 | 0 | 0 | 0 | 0 | 0.0003 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | |
| з | 0.0022 | 0.0018 | 0.0001 | 0.0004 | 0.0001 | 0.0006 | 0.0001 | 0 | 0 | 0 | 0 | 0 | 0.0005 | 0.0011 | 0.0026 | 0.0011 | 0 | 0.0002 | 0 | 0 | 0.0003 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.0009 | 0 | 0 | 0.0002 | |
| и | 0.0216 | 0.0001 | 0.0016 | 0.002 | 0.0002 | 0.0014 | 0.0039 | 0.0009 | 0.0029 | 0.0011 | 0.0006 | 0.0011 | 0.0026 | 0.0062 | 0.0043 | 0.0001 | 0 | 0.0006 | 0.0026 | 0.0037 | 0 | 0.0001 | 0.0022 | 0.0006 | 0.0029 | 0.0005 | 0.0001 | 0 | 0 | 0 | 0 | 0.0005 | 0.0025 | |
| й | 0.0042 | 0 | 0 | 0 | 0 | 0.0003 | 0 | 0 | 0 | 0 | 0 | 0 | 0.0002 | 0 | 0 | 0.0002 | 0 | 0 | 0.0007 | 0 | 0 | 0 | 0 | 0.0001 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | |
| к | 0.0053 | 0.0067 | 0 | 0.0001 | 0 | 0.0004 | 0.0003 | 0.0013 | 0 | 0.0003 | 0 | 0 | 0.0003 | 0.0004 | 0.0001 | 0.0001 | 0 | 0.0006 | 0.0001 | 0.0006 | 0.0008 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | |
| л | 0.0008 | 0.0026 | 0 | 0 | 0 | 0 | 0.0005 | 0.0005 | 0.0075 | 0 | 0 | 0 | 0 | 0.0001 | 0.005 | 0 | 0 | 0 | 0.0001 | 0.0012 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.0001 | 0.0041 | 0 | 0.0003 | 0.0015 | |
| м | 0.0097 | 0.0024 | 0 | 0 | 0 | 0 | 0.0056 | 0 | 0.0018 | 0 | 0 | 0.0001 | 0 | 0.0008 | 0.0044 | 0.0001 | 0 | 0.0001 | 0 | 0.0028 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.0003 | |
| н | 0.0009 | 0.0083 | 0 | 0 | 0 | 0 | 0.0101 | 0 | 0 | 0.0115 | 0 | 0 | 0 | 0 | 0.0032 | 0.017 | 0 | 0 | 0.0002 | 0.0003 | 0.0004 | 0 | 0 | 0.0001 | 0.0001 | 0 | 0 | 0 | 0.0034 | 0.0003 | 0 | 0 | 0.0004 | |
| о | 0.0389 | 0 | 0.0049 | 0.0048 | 0.0048 | 0.0073 | 0.0047 | 0.0032 | 0.0022 | 0.0009 | 0.0025 | 0.0008 | 0.0042 | 0.0069 | 0.0034 | 0.0006 | 0.0019 | 0.0037 | 0.0084 | 0.0063 | 0 | 0.0002 | 0.0001 | 0 | 0.0012 | 0.0006 | 0 | 0 | 0 | 0 | 0.0004 | 0 | 0.0006 | |
| п | 0 | 0.0002 | 0 | 0 | 0 | 0 | 0.0012 | 0 | 0 | 0.0002 | 0 | 0 | 0.0004 | 0 | 0 | 0.0071 | 0 | 0.0091 | 0 | 0 | 0.0002 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.0001 | 0 | 0.0001 | |
| р | 0.0009 | 0.0052 | 0 | 0.0009 | 0 | 0 | 0.0045 | 0.0006 | 0 | 0.0061 | 0 | 0 | 0 | 0.0004 | 0.0002 | 0.0048 | 0.0001 | 0 | 0 | 0.0001 | 0.0027 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.0009 | 0.0001 | 0 | 0 | 0.0006 | |
| с | 0.0012 | 0.0016 | 0 | 0.0008 | 0 | 0 | 0.0026 | 0 | 0 | 0.0006 | 0 | 0.0022 | 0.005 | 0.0009 | 0.001 | 0.003 | 0.0013 | 0.0002 | 0.0008 | 0.0141 | 0.0043 | 0 | 0.0003 | 0 | 0.0004 | 0 | 0 | 0 | 0.0002 | 0.0005 | 0 | 0.0001 | 0.0041 | |
| т | 0.0091 | 0.0057 | 0 | 0.0051 | 0 | 0.0004 | 0.0045 | 0 | 0.0046 | 0.0002 | 0.0002 | 0.0003 | 0.0004 | 0.0001 | 0.0013 | 0.0194 | 0 | 0.0016 | 0.003 | 0 | 0.0002 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.0004 | 0.0076 | 0 | 0 | 0.0004 | |
| у | 0.0054 | 0 | 0.0002 | 0.0005 | 0.0024 | 0.0016 | 0.0007 | 0.0006 | 0 | 0 | 0.0001 | 0.0007 | 0 | 0.0004 | 0.0002 | 0 | 0.0002 | 0.0001 | 0.0005 | 0.0016 | 0 | 0 | 0.0002 | 0 | 0.0012 | 0.0002 | 0.0037 | 0 | 0 | 0 | 0 | 0.001 | 0 | |
| ф | 0 | 0.0001 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.0002 | 0 | 0 | 0 | 0 | 0 | 0.0003 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| х | 0.0045 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.0022 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | |
| ц | 0.0001 | 0.0004 | 0 | 0.0001 | 0 | 0 | 0.0004 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.0002 | 0 | 0 | 0 | |
| ч | 0 | 0.0029 | 0 | 0 | 0 | 0 | 0.0048 | 0 | 0 | 0.0031 | 0 | 0.0001 | 0 | 0 | 0.0012 | 0 | 0 | 0 | 0.0058 | 0.0004 | 0 | 0 | 0 | 0 | 0 | 0 | 0.0001 | 0 | 0 | 0 | 0.0001 | 0 | 0 | 0 |
| ш | 0 | 0.0003 | 0 | 0 | 0 | 0 | 0.0018 | 0 | 0 | 0.0003 | 0 | 0 | 0.0001 | 0 | 0 | 0.0001 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.0002 | 0 | 0 | 0 | |
| щ | 0 | 0.0003 | 0 | 0 | 0 | 0 | 0.0044 | 0 | 0 | 0.0015 | 0 | 0 | 0 | 0 | 0.0012 | 0 | 0 | 0 | 0 | 0 | 0 | 0.0001 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.0001 | 0 | 0 | 0 | 0 |
| ъ | 0 | 0 | 0 | 0 | 0 | 0 | 0.0001 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.0001 | |
| ы | 0.004 | 0 | 0 | 0.0013 | 0 | 0 | 0.0011 | 0 | 0.0001 | 0 | 0.0004 | 0 | 0.0009 | 0.0017 | 0 | 0 | 0 | 0.0001 | 0.001 | 0.0015 | 0 | 0 | 0.0011 | 0 | 0 | 0.0002 | 0 | 0 | 0 | 0 | 0 | 0 | 0.0001 | |
| ь | 0.0092 | 0 | 0 | 0 | 0 | 0 | 0.0001 | 0 | 0.0002 | 0 | 0 | 0.0009 | 0 | 0.0001 | 0.0019 | 0 | 0 | 0 | 0.0006 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.0004 | 0.0001 | |
| э | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.0003 | 0 | 0.0003 | 0 | 0 | 0 | 0 | 0 | 0 | 0.0033 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| ю | 0.0014 | 0 | 0.0001 | 0 | 0 | 0.0001 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.0017 | 0 | 0 | 0 | 0 | 0.0001 | 0 | 0.001 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| я | 0.0091 | 0 | 0 | 0.0002 | 0 | 0.0002 | 0.0003 | 0.0001 | 0.0001 | 0.0001 | 0 | 0 | 0.0003 | 0.0001 | 0.0004 | 0.0003 | 0 | 0 | 0.0005 | 0.0013 | 0 | 0 | 0.0004 | 0 | 0 | 0 | 0 | 0.0007 | 0 | 0 | 0 | 0 | 0.0002 | |

Рисунок 3: Біграми з пробілами, з перетинами

| | а | б | в | г | д | е | ж | з | и | й | к | л | м | н | о | п | р | с | т | у | ф | х | ц | ч | ш | щ | ъ | ы | ь | э | ю | я | | | |
|---|--------|---------|---------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|
| а | 0 | 0.0032 | 0.0059 | 0.0143 | 0.0014 | 0.0083 | 0.0057 | 0.0028 | 0.0022 | 0.0177 | 0 | 0.0073 | 0.002 | 0.0049 | 0.0158 | 0.0142 | 0.0135 | 0.0031 | 0.0139 | 0.0124 | 0.0026 | 0.0005 | 0.0004 | 0.0004 | 0.0091 | 0.0001 | 0 | 0 | 0 | 0 | 0.0036 | 0 | 0.0006 | | |
| б | 0.0003 | 0.0002 | 0 | 0 | 0 | 0.0018 | 0.0006 | 0.0009 | 0.0033 | 0.0009 | 0.0042 | 0.0002 | 0.0002 | 0.0073 | 0.0028 | 0.0021 | 0.0024 | 0 | 0.0009 | 0.0003 | 0.0017 | 0.0039 | 0.0004 | 0 | 0.0006 | 0 | 0.0022 | 0.0001 | 0.0003 | 0 | 0 | 0 | 0.0014 | 0.0015 | |
| в | 0.0052 | 0.0029 | 0 | 0 | 0 | 0 | 0.0001 | 0.0057 | 0 | 0.0001 | 0.0029 | 0 | 0 | 0.0008 | 0.0002 | 0.0008 | 0.0082 | 0.0001 | 0.0004 | 0.0025 | 0.0001 | 0.0014 | 0 | 0.0004 | 0 | 0 | 0.0001 | 0 | 0 | 0.0012 | 0.0001 | 0 | 0 | 0.0001 | |
| г | 0.0003 | 0.0007 | 0 | 0 | 0 | 0.0009 | 0.0001 | 0 | 0 | 0.0008 | 0 | 0.0001 | 0.0006 | 0 | 0.0001 | 0.0083 | 0 | 0.0002 | 0 | 0 | 0.0003 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.0001 | |
| д | 0.0004 | 0.0029 | 0 | 0.0016 | 0 | 0 | 0.0059 | 0 | 0 | 0.0032 | 0 | 0.0001 | 0.0011 | 0.0003 | 0.0035 | 0.0032 | 0.0001 | 0.0023 | 0.0002 | 0.0001 | 0.0019 | 0 | 0 | 0.0001 | 0 | 0.0002 | 0 | 0 | 0.0004 | 0.0011 | 0 | 0 | 0.0008 | | |
| е | 0.0249 | 0 | 0.0008 | 0.0008 | 0.0024 | 0.0056 | 0.003 | 0.0012 | 0.0005 | 0.0001 | 0.0018 | 0.0015 | 0.0059 | 0.0045 | 0.0084 | 0.0007 | 0.0005 | 0.0047 | 0.0075 | 0.0072 | 0 | 0 | 0.0006 | 0.0001 | 0.0013 | 0.0001 | 0.0012 | 0 | 0 | 0 | 0 | 0.0004 | 0.0002 | | |
| ж | 0 | 0.0004 | 0 | 0 | 0 | 0.0019 | 0.0056 | 0 | 0 | 0.0006 | 0 | 0 | 0 | 0 | 0.0021 | 0 | 0 | 0 | 0 | 0 | 0.0003 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | |
| з | 0.0021 | 0.0019 | 0 | 0.0004 | 0.0001 | 0.0006 | 0.0001 | 0 | 0.0001 | 0 | 0 | 0 | 0.0005 | 0.0011 | 0.0025 | 0.001 | 0 | 0.0002 | 0 | 0 | 0.0003 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.0009 | 0 | 0 | 0.0002 | | |
| и | 0.0212 | 0.0001 | 0.0016 | 0.002 | 0.0001 | 0.0014 | 0.0038 | 0.0009 | 0.003 | 0.0011 | 0.0006 | 0.0011 | 0.0026 | 0.0064 | 0.0043 | 0.0001 | 0 | 0.0006 | 0.0027 | 0.0038 | 0 | 0.0001 | 0.0023 | 0.0005 | 0.0029 | 0.0005 | 0.0001 | 0 | 0 | 0 | 0 | 0.0005 | 0.0024 | | |
| й | 0.0042 | 0 | 0 | 0 | 0 | 0.0003 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.0002 | 0 | 0 | 0.0002 | 0 | 0 | 0.0007 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | |
| к | 0.0054 | 0.0068 | 0 | 0 | 0 | 0 | 0.0001 | 0.0003 | 0 | 0.0015 | 0 | 0 | 0.0003 | 0 | 0.0002 | 0.0063 | 0 | 0.0007 | 0.0001 | 0.0006 | 0.0008 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | |
| л | 0.0007 | 0.0027 | 0 | 0 | 0 | 0.005 | 0.0006 | 0 | 0 | 0.0016 | 0 | 0.0001 | 0.0049 | 0 | 0.0012 | 0.0002 | 0 | 0 | 0 | 0.0001 | 0.0012 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.0001 | 0.004 | 0 | 0.0003 | 0.0016 | | |
| м | 0.0093 | 0.0023 | 0 | 0 | 0 | 0 | 0.0057 | 0 | 0 | 0.0018 | 0 | 0 | 0.0001 | 0.0008 | 0.0043 | 0.0001 | 0 | 0.0001 | 0.0002 | 0.0029 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.0021 | 0 | 0 | 0.0003 | 0.0003 | |
| н | 0.0009 | 0.0083 | 0 | 0 | 0 | 0 | 0.0101 | 0 | 0 | 0.0116 | 0 | 0 | 0 | 0.0032 | 0.0172 | 0 | 0 | 0.0002 | 0.0004 | 0.0005 | 0 | 0 | 0.0001 | 0.0001 | 0 | 0 | 0 | 0 | 0.0033 | 0.0004 | 0 | 0 | 0.0004 | 0.0004 | |
| о | 0.0337 | 0 | 0.0049 | 0.0048 | 0.0046 | 0.0073 | 0.0047 | 0.0047 | 0.0033 | 0.0029 | 0.0009 | 0.0025 | 0.0008 | 0.0042 | 0.007 | 0.0034 | 0.0005 | 0.0019 | 0.0037 | 0.0086 | 0.0063 | 0 | 0.0002 | 0.0001 | 0 | 0.0011 | 0.0006 | 0 | 0 | 0 | 0.0004 | 0 | 0.0005 | 0.0005 | |
| п | 0 | 0.0002 | 0 | 0 | 0 | 0 | 0.0013 | 0 | 0 | 0.0002 | 0 | 0 | 0.0004 | 0 | 0 | 0.0069 | 0 | 0.0092 | 0 | 0 | 0.0002 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.0001 | 0 | 0 | 0.0001 | 0.0001 | |
| р | 0.001 | 0.0051 | 0 | 0.0009 | 0 | 0.0045 | 0.0007 | 0 | 0.0061 | 0 | 0 | 0 | 0.0004 | 0.0002 | 0.0046 | 0.0001 | 0 | 0.0001 | 0.0001 | 0.0027 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.0001 | 0 | 0.0009 | 0.0001 | 0 | 0.0006 | 0.0006 | |
| с | 0.0012 | 0.0016 | 0 | 0.0008 | 0 | 0 | 0.0027 | 0 | 0.0006 | 0 | 0.0021 | 0.0049 | 0.0008 | 0.001 | 0.0029 | 0.0013 | 0.0002 | 0.0008 | 0.0138 | 0.0044 | 0.0004 | 0.0004 | 0 | 0.0004 | 0 | 0.0004 | 0 | 0 | 0.0002 | 0.0005 | 0 | 0.0001 | 0.0004 | 0.0004 | |
| т | 0 | 0.00056 | 0 | 0.0053 | 0 | 0.0004 | 0.0044 | 0 | 0.0046 | 0 | 0.0002 | 0.0003 | 0 | 0.0013 | 0.0094 | 0 | 0.0016 | 0.0031 | 0 | 0.0001 | 0 | 0 | 0 | 0.0001 | 0.0024 | 0 | 0 | 0.0004 | 0.0075 | 0 | 0 | 0 | 0.0005 | 0.0005 | |
| у | 0.0053 | 0.0002 | 0.0002 | 0.0024 | 0.0015 | 0.0008 | 0.0016 | 0.0001 | 0.0001 | 0.0002 | 0 | 0.0001 | 0.0005 | 0.0002 | 0 | 0.0004 | 0.0002 | 0.0002 | 0.0005 | 0.0015 | 0 | 0.0001 | 0.0001 | 0 | 0.0011 | 0.0004 | 0.0001 | 0 | 0 | 0 | 0.0001 | 0.0001 | 0.0001 | | |
| ф | 0 | 0 | 0 | 0 | 0 | 0.0001 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.0021 | 0 | 0 | 0 | 0 | 0.0001 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | |
| х | 0.0044 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.0011 | 0 | 0 | 0 | 0 | 0.0001 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | |
| ц | 0.0001 | 0.0004 | 0.0001 | 0 | 0 | 0.0004 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.0002 | 0 | 0 | 0 | 0 | |
| ч | 0 | 0.0028 | 0 | 0 | 0 | 0.0048 | 0 | 0 | 0.0003 | 0 | 0.0001 | 0 | 0 | 0.0012 | 0 | 0 | 0 | 0 | 0.0057 | 0.0005 | 0 | 0 | 0 | 0 | 0 | 0.0001 | 0 | 0 | 0 | 0 | 0.0001 | 0 | 0 | 0 | 0 |
| ш | 0 | 0.0003 | 0 | 0 | 0 | 0.0018 | 0 | 0 | 0.0003 | 0 | 0.0001 | 0 | 0 | 0.0001 | 0 | 0.0001 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.0002 | 0 | 0 | 0 | 0 |
| щ | 0 | 0.0003 | 0 | 0 | 0 | 0.0044 | 0 | 0.0016 | 0 | 0 | 0 | 0 | 0 | 0.0013 | 0 | 0 | 0 | 0 | 0 | 0 | 0.0001 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.0002 | 0 | 0 | 0 | 0 |
| ъ | 0 | 0 | 0 | 0 | 0 | 0 | 0.0001 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.0001 | |
| ы | 0.0093 | 0 | 0 | 0.0011 | 0 | 0.0001 | 0 | 0.0004 | 0 | 0.0001 | 0.0004 | 0 | 0.0019 | 0 | 0.001 | 0.0001 | 0.0001 | 0.0001 | 0.0001 | 0.0001 | 0.0001 | 0.0001 | 0.0001 | 0.0001 | 0.0001 | 0.0001 | 0.0001 | 0.0001 | 0.0001 | 0.0001 | 0.0001 | 0.0001 | 0.0001 | 0.0001 | |
| ь | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.0002 | 0 | 0.0009 | 0 | 0.0001 | 0.0019 | 0 | 0 | 0 | 0.0006 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.0004 | 0.0001 | 0.0001 |
| э | 0 | 0.0014 | 0.00001 | 0 | 0.0001 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.0017 | 0 | 0 | 0 | 0 | 0 | 0.0001 | 0 | 0.0001 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| ю | 0.0091 | 0 | 0 | 0.0002 | 0.0002 | 0.0004 | 0.0001 | 0 | 0 | 0.0003 | 0.0001 | 0.0004 | 0.0003 | 0 | 0 | 0 | 0 | 0.0005 | 0.0013 | 0 | 0 | 0.0004 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.0007 | 0 | 0 | 0.0002 | 0.0001 | 0.0001 |

| | а | б | в | г | д | е | ж | з | и | й | к | л | м | н | о | п | р | с | т | у | ф | х | ц | ч | ш | щ | ъ | ы | ь | э | ю | я | |
|---|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|
| а | 0.0003 | 0.0004 | 0.0033 | 0.0008 | 0.0017 | 0.0045 | 0.0012 | 0.0051 | 0.0019 | 0.0002 | 0.0092 | 0.0034 | 0.0028 | 0.0041 | 0.0009 | 0.0021 | 0.0006 | 0.0029 | 0.0056 | 0.0007 | 0 | 0.0007 | 0 | 0.003 | 0.0001 | 0.0003 | 0 | 0 | 0 | 0.0003 | 0.0017 | 0.0019 | |
| б | 0.0002 | 0 | 0 | 0 | 0 | 0.0015 | 0 | 0 | 0.0002 | 0 | 0 | 0.0008 | 0 | 0.0009 | 0.0034 | 0 | 0.0015 | 0.0005 | 0 | 0.0013 | 0 | 0.0006 | 0 | 0 | 0 | 0 | 0.0007 | 0.0002 | 0.0044 | 0 | 0.0001 | 0 | 0.0001 |
| в | 0.0046 | 0.0002 | 0.0005 | 0 | 0.0008 | 0.0069 | 0 | 0.0002 | 0.0038 | 0 | 0.0004 | 0.001 | 0.0003 | 0.0014 | 0.0107 | 0.0005 | 0.0006 | 0.0041 | 0.0008 | 0.0018 | 0 | 0.0005 | 0 | 0.0002 | 0.0001 | 0 | 0 | 0.0015 | 0.0001 | 0.0002 | 0 | 0.0002 | |
| г | 0.0008 | 0 | 0 | 0 | 0.0011 | 0.0001 | 0 | 0 | 0.0009 | 0 | 0.0001 | 0.0008 | 0 | 0.0001 | 0.0098 | 0 | 0.0002 | 0 | 0 | 0.0004 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| д | 0.0035 | 0 | 0.002 | 0 | 0 | 0.007 | 0 | 0 | 0.0039 | 0 | 0.0001 | 0.0014 | 0.0004 | 0.0043 | 0.0039 | 0.0002 | 0.0028 | 0.0003 | 0.0001 | 0.0022 | 0 | 0 | 0.0001 | 0 | 0.0002 | 0 | 0 | 0 | 0.0005 | 0.0014 | 0 | 0 | 0.001 |
| е | 0.0008 | 0.0021 | 0.0037 | 0.0033 | 0.0079 | 0.0049 | 0.0018 | 0.001 | 0.0036 | 0.0022 | 0.0032 | 0.0075 | 0.0067 | 0.0126 | 0.0031 | 0.0032 | 0.006 | 0.0117 | 0.0107 | 0.0005 | 0.0001 | 0.0006 | 0.0002 | 0.0029 | 0.0002 | 0.0016 | 0 | 0 | 0 | 0.0005 | 0.0005 | 0.0003 | |
| ж | 0.0006 | 0 | 0 | 0 | 0.0022 | 0.0066 | 0 | 0 | 0.0008 | 0 | 0 | 0 | 0 | 0.0026 | 0 | 0 | 0 | 0 | 0 | 0.0004 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| з | 0.0022 | 0.0001 | 0.0006 | 0.0001 | 0.0009 | 0.0002 | 0 | 0.0001 | 0.0001 | 0 | 0.0002 | 0.0006 | 0.0014 | 0.0036 | 0.0015 | 0.0002 | 0.0003 | 0.0003 | 0.0002 | 0.0004 | 0 | 0 | 0 | 0.0002 | 0 | 0 | 0 | 0.0001 | 0.0011 | 0 | 0.0002 | 0 | 0.0002 |
| и | 0.0005 | 0.0029 | 0.0046 | 0.0005 | 0.0029 | 0.0054 | 0.0015 | 0.0037 | 0.0035 | 0.0007 | 0.0025 | 0.0034 | 0.0082 | 0.0078 | 0.0023 | 0.0024 | 0.0012 | 0.0052 | 0.0071 | 0.0005 | 0.0002 | 0.0028 | 0.0007 | 0.0044 | 0.0006 | 0.0001 | 0 | 0 | 0 | 0.0006 | 0.0007 | 0.003 | |
| й | 0.0002 | 0.0001 | 0.0004 | 0 | 0.0006 | 0.0001 | 0.0001 | 0 | 0.0007 | 0 | 0.0004 | 0.0001 | 0.0004 | 0.0007 | 0.0002 | 0.0005 | 0.0001 | 0.0015 | 0.0003 | 0.0001 | 0 | 0 | 0.0001 | 0.0002 | 0.0001 | 0 | 0 | 0 | 0 | 0.0001 | 0 | 0 | |
| к | 0.008 | 0.0002 | 0.0004 | 0.0001 | 0.0002 | 0.0003 | 0.0009 | 0 | 0.0023 | 0 | 0.0005 | 0.0004 | 0.0003 | 0.0008 | 0.0082 | 0.0004 | 0.0008 | 0.0005 | 0.0011 | 0.0011 | 0 | 0 | 0 | 0.0005 | 0 | 0 | 0 | 0 | 0 | 0.0002 | 0 | 0 | |
| л | 0.0032 | 0.0001 | 0 | 0 | 0 | 0.006 | 0.0006 | 0 | 0.0091 | 0 | 0.0001 | 0.0001 | 0 | 0.0002 | 0.0061 | 0.0001 | 0 | 0.0001 | 0.0001 | 0.0014 | 0 | 0 | 0 | 0.0001 | 0 | 0 | 0 | 0.0001 | 0.0049 | 0 | 0.0003 | 0.0018 | |
| м | 0.0032 | 0.0002 | 0.0007 | 0.0002 | 0.0007 | 0.007 | 0.0004 | 0.0001 | 0.0036 | 0 | 0.0004 | 0.0003 | 0.0003 | 0.0019 | 0.0067 | 0.0008 | 0.0002 | 0.0014 | 0.0007 | 0.0035 | 0 | 0 | 0 | 0.0007 | 0 | 0 | 0 | 0.0024 | 0 | 0.0001 | 0 | 0.0004 | |
| н | 0.01 | 0.0001 | 0.0001 | 0 | 0.0001 | 0.0121 | 0 | 0 | 0.014 | 0 | 0 | 0 | 0 | 0.0039 | 0.0204 | 0.0001 | 0 | 0.0003 | 0.0005 | 0.0006 | 0 | 0 | 0.0001 | 0.0001 | 0 | 0 | 0 | 0 | 0.0041 | 0.0004 | 0 | 0 | 0.0005 |
| о | 0.0005 | 0.0078 | 0.0095 | 0.0059 | 0.0105 | 0.007 | 0.0047 | 0.003 | 0.0053 | 0.003 | 0.0023 | 0.0055 | 0.0091 | 0.0078 | 0.0043 | 0.0056 | 0.0053 | 0.0136 | 0.0105 | 0.0007 | 0.0003 | 0.0002 | 0.0001 | 0.0042 | 0.0008 | 0 | 0 | 0 | 0 | 0.0011 | 0 | 0.0009 | |
| п | 0.0002 | 0 | 0 | 0 | 0 | 0.0015 | 0 | 0 | 0.0002 | 0 | 0 | 0.0008 | 0 | 0.0001 | 0.0085 | 0 | 0.0109 | 0 | 0 | 0.0002 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.0001 | 0 | 0 | 0.0001 | |
| р | 0.0063 | 0 | 0.0012 | 0 | 0.0001 | 0.0055 | 0.0008 | 0 | 0.0074 | 0 | 0.0001 | 0 | 0.0005 | 0.0003 | 0.0058 | 0.0002 | 0 | 0.0001 | 0.0002 | 0.0032 | 0 | 0 | 0 | 0 | 0.0001 | 0.0001 | 0 | 0 | 0.0011 | 0.0001 | 0 | 0 | 0.0007 |
| с | 0.0019 | 0 | 0.001 | 0 | 0.0003 | 0.0032 | 0 | 0 | 0.0008 | 0 | 0.0027 | 0.006 | 0.0011 | 0.0013 | 0.0037 | 0.0017 | 0.0003 | 0.001 | 0.017 | 0.0052 | 0 | 0.0004 | 0 | 0.0005 | 0 | 0 | 0 | 0 | 0.0003 | 0.0006 | 0 | 0.0001 | 0.0049 |
| т | 0.0071 | 0.0005 | 0.0069 | 0.0001 | 0.0011 | 0.0059 | 0.0001 | 0.0002 | 0.0067 | 0 | 0.0007 | 0.0006 | 0.0003 | 0.0027 | 0.0241 | 0.0009 | 0.0021 | 0.0043 | 0.0006 | 0.0004 | 0 | 0 | 0 | 0.0008 | 0 | 0 | 0 | 0 | 0.0005 | 0.0091 | 0.0003 | 0 | 0.0006 |
| у | 0.0002 | 0.0004 | 0.001 | 0.0029 | 0.0021 | 0.001 | 0.0009 | 0.0001 | 0.0008 | 0.0001 | 0.0012 | 0.0001 | 0.0006 | 0.001 | 0.0005 | 0.0007 | 0.0003 | 0.001 | 0.0026 | 0.0001 | 0 | 0.0002 | 0 | 0.0021 | 0.0002 | 0.0044 | 0 | 0 | 0 | 0.0001 | 0.0012 | 0 | |
| ф | 0.0001 | 0 | 0 | 0 | 0 | 0.0001 | 0 | 0 | 0.0003 | 0 | 0 | 0 | 0 | 0 | 0.0004 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| х | 0.0001 | 0.0001 | 0.0006 | 0.0001 | 0.0002 | 0.0001 | 0.0001 | 0.0001 | 0.0005 | 0 | 0.0003 | 0.0001 | 0.0002 | 0.0005 | 0.003 | 0.0005 | 0.0002 | 0.0007 | 0.0003 | 0.0001 | 0 | 0 | 0 | 0.0002 | 0 | 0 | 0 | 0 | 0 | 0.0001 | 0 | 0 | |
| ц | 0.0005 | 0 | 0.0001 | 0 | 0 | 0.0005 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.0002 | 0 | 0 | 0 | |
| ч | 0.0034 | 0 | 0 | 0 | 0 | 0.0057 | 0 | 0 | 0.0037 | 0 | 0.0001 | 0 | 0 | 0.0014 | 0 | 0 | 0 | 0 | 0.0069 | 0.0005 | 0 | 0 | 0 | 0 | 0.0001 | 0 | 0 | 0 | 0 | 0.0001 | 0 | 0 | 0 |
| ш | 0.0004 | 0 | 0 | 0 | 0.0022 | 0 | 0 | 0.0004 | 0 | 0 | 0.0002 | 0 | 0 | 0.0001 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.0003 | 0 | 0 | 0 |
| щ | 0.0003 | 0 | 0 | 0 | 0.0052 | 0 | 0 | 0.0018 | 0 | 0 | 0 | 0 | 0 | 0.0015 | 0 | 0 | 0 | 0 | 0 | 0.0001 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.0002 | 0 | 0 | 0 |
| ъ | 0 | 0 | 0 | 0 | 0 | 0.0001 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.0001 | |
| ы | 0.0001 | 0.0003 | 0.0019 | 0.0001 | 0.0003 | 0.0014 | 0 | 0.0002 | 0.0006 | 0.0005 | 0.0002 | 0.0012 | 0.0021 | 0.0006 | 0.0004 | 0.0005 | 0.0002 | 0.0016 | 0.0021 | 0.0001 | 0 | 0.0014 | 0 | 0.0001 | 0.0002 | 0 | 0 | 0 | 0 | 0.0001 | 0 | 0.0001 | |
| ь | 0.0002 | 0.0005 | 0.001 | 0 | 0.0005 | 0.0007 | 0.0001 | 0.0004 | 0.0015 | 0 | 0.0016 | 0.0002 | 0.0004 | 0.0037 | 0.0008 | 0.0009 | 0.0001 | 0.0015 | 0.0007 | 0.0001 | 0 | 0 | 0 | 0.0007 | 0.0009 | 0 | 0 | 0 | 0 | 0.0002 | 0.0005 | 0.0001 | |
| э | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.0003 | 0 | 0.0004 | 0.0001 | 0 | 0 | 0 | 0 | 0 | 0 | 0.0039 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| ю | 0.0001 | 0.0001 | 0.0002 | 0 | 0.0002 | 0 | 0 | 0 | 0.0002 | 0 | 0.0001 | 0 | 0.0001 | 0.0001 | 0.0001 | 0.0002 | 0 | 0.0001 | 0.0021 | 0 | 0 | 0 | 0 | 0 | 0.0002 | 0 | 0.0012 | 0 | 0 | 0 | 0 | 0 | 0 |
| я | 0.0003 | 0.0003 | 0.0017 | 0.0001 | 0.0007 | 0.0007 | 0.0001 | 0.0003 | 0.0014 | 0 | 0.0011 | 0.0003 | 0.0008 | 0.0014 | 0.0008 | 0.0008 | 0.0002 | 0.0013 | 0.0024 | 0.0002 | 0 | 0.0005 | 0 | 0.0006 | 0 | 0.0009 | 0 | 0 | 0 | 0.0002 | 0.0002 | 0.0001 | |

Рисунок 5: Біграми без пробілів, з перетинами

| а | б | в | г | д | е | ж | з | и | й | к | л | м | н | о | п | р | с | т | у | ф | х | ц | ч | ш | щ | ъ | ы | ь | э | ю | я | | | |
|---|--------|--------|--------|--------|--------|--------|--------|--------|----------|--------|----------|----------|----------|----------|----------|----------|----------|----------|----------|----------|----------|----------|----------|----------|----------|----------|----------|----------|----------|----------|----------|----------|----------|--------|
| а | 0.0003 | 0.0004 | 0.0032 | 0.0008 | 0.0017 | 0.0045 | 0.0011 | 0.0052 | 0.0019 | 0.0003 | 0.0094 | 0.0034 | 0.0028 | 0.0041 | 0.0009 | 0.0022 | 0.0006 | 0.0028 | 0.0056 | 0.0006 | 0 | 0.0007 | 0 | 0.003 | 0.0002 | 0.0003 | 0 | 0 | 0 | 0.0003 | 0.0017 | 0.002 | | |
| б | 0.0003 | 0 | 0 | 0 | 0.0016 | 0 | 0 | 0.0002 | 0 | 0 | 0.0008 | 0 | 0.001 | 0.0034 | 0 | 0.0015 | 0.0005 | 0 | 0.0013 | 0 | 0.0005 | 0 | 0 | 0 | 0.0007 | 0.0002 | 0.0043 | 0 | 0.0001 | 0 | 0.0001 | 0 | 0.0001 | |
| в | 0.0045 | 0.0002 | 0.0005 | 0 | 0.0008 | 0.007 | 0 | 0.0002 | 0.0038 | 0 | 0.0003 | 0.001 | 0.0003 | 0.0014 | 0.0105 | 0.0004 | 0.0006 | 0.0043 | 0.0008 | 0.0018 | 0 | 0.0005 | 0 | 0.0002 | 0.0001 | 0 | 0 | 0.0014 | 0.0001 | 0.0002 | 0 | 0.0002 | | |
| г | 0.0007 | 0 | 0 | 0 | 0.001 | 0.0002 | 0 | 0 | 0.0009 | 0 | 0.0001 | 0.0008 | 0 | 0.0001 | 0.0098 | 0 | 0.0002 | 0 | 0 | 0.0004 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | |
| д | 0.0036 | 0 | 0.002 | 0 | 0 | 0.007 | 0 | 0 | 0.0038 | 0 | 0.0001 | 0.0014 | 0.0004 | 0.0042 | 0.0039 | 0.0002 | 0.0028 | 0.0003 | 0.0001 | 0.0022 | 0 | 0 | 0.0001 | 0 | 0.0002 | 0 | 0 | 0.0005 | 0.0014 | 0 | 0 | 0.001 | | |
| е | 0.0007 | 0.0021 | 0.0038 | 0.0033 | 0.0082 | 0.0049 | 0.0019 | 0.0011 | 0.0035 | 0.0023 | 0.0032 | 0.0076 | 0.0066 | 0.0126 | 0.003 | 0.0031 | 0.0058 | 0.0118 | 0.0104 | 0.0005 | 0 | 0.0008 | 0.0002 | 0.0028 | 0.0002 | 0.0015 | 0 | 0 | 0 | 0.0006 | 0.0005 | 0.0003 | | |
| ж | 0.0005 | 0 | 0 | 0 | 0.0021 | 0.0065 | 0 | 0 | 0.0008 | 0 | 0 | 0 | 0 | 0.0027 | 0 | 0 | 0 | 0 | 0.0004 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | |
| з | 0.0022 | 0.0001 | 0.0007 | 0.0001 | 0.0009 | 0.0002 | 0 | 0.0001 | 0.0001 | 0 | 0.0002 | 0.0006 | 0.0013 | 0.0036 | 0.0015 | 0.0002 | 0.0002 | 0.0003 | 0.0002 | 0.0004 | 0 | 0 | 0 | 0.0002 | 0 | 0 | 0.0001 | 0.0001 | 0 | 0.0002 | 0 | 0.0002 | | |
| и | 0.0006 | 0.0029 | 0.0047 | 0.0005 | 0.0029 | 0.0055 | 0.0014 | 0.0038 | 0.0034 | 0.0007 | 0.0025 | 0.0034 | 0.0098 | 0.0008 | 0.0023 | 0.0024 | 0.0012 | 0.0051 | 0.0074 | 0.0006 | 0.0002 | 0.0027 | 0.0008 | 0.0044 | 0.0006 | 0.0001 | 0 | 0 | 0 | 0.0006 | 0.0007 | 0.0029 | | |
| й | 0.0002 | 0.0001 | 0.0004 | 0 | 0.0005 | 0.0001 | 0.0001 | 0 | 0.0007 | 0 | 0.0004 | 0.0006 | 0.0002 | 0.0004 | 0.0001 | 0.0015 | 0.0003 | 0.0003 | 0.0001 | 0 | 0 | 0.0001 | 0.0002 | 0.0001 | 0 | 0 | 0 | 0.0001 | 0 | 0 | 0 | 0 | 0 | |
| к | 0.0079 | 0.0002 | 0.0004 | 0.0005 | 0.0002 | 0.0001 | 0 | 0 | 0.0002 | 0.0002 | 0.0001 | 0.0001 | 0.0003 | 0.0009 | 0.0005 | 0.0012 | 0.0001 | 0 | 0.0006 | 0.0004 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.0002 | 0 | 0 | 0 | 0 | |
| л | 0.0032 | 0 | 0.0001 | 0 | 0 | 0.0058 | 0.0006 | 0 | 0.0001 | 0 | 0.0001 | 0 | 0 | 0.0002 | 0.006 | 0.0001 | 0 | 0.0001 | 0.0001 | 0.0014 | 0 | 0 | 0.0001 | 0 | 0 | 0 | 0 | 0.0002 | 0.0049 | 0 | 0.0003 | 0.0018 | 0.008 | |
| м | 0.0033 | 0.0002 | 0.0007 | 0.0002 | 0.0007 | 0.007 | 0.0004 | 0.0001 | 0.0036 | 0 | 0.0004 | 0.0003 | 0.0003 | 0.002 | 0.0067 | 0.0008 | 0.0002 | 0.014 | 0.0007 | 0.0036 | 0 | 0 | 0.0007 | 0 | 0 | 0 | 0.0024 | 0 | 0 | 0 | 0.0001 | 0.0005 | | |
| н | 0.01 | 0.0001 | 0.0001 | 0 | 0.0001 | 0.0121 | 0 | 0 | 0.014 | 0 | 0 | 0.0039 | 0.0205 | 0.0001 | 0 | 0.0036 | 0.0004 | 0.0005 | 0 | 0 | 0 | 0.0001 | 0.0002 | 0 | 0 | 0 | 0.0041 | 0.0004 | 0 | 0 | 0.0005 | 0 | 0.0005 | |
| о | 0.006 | 0.0079 | 0.0096 | 0.006 | 0.0103 | 0.0071 | 0.0048 | 0.0032 | 0.0053 | 0.0032 | 0.0022 | 0.0053 | 0.0091 | 0.0078 | 0.0043 | 0.0057 | 0.0054 | 0.0136 | 0.0105 | 0.0007 | 0.0003 | 0.0002 | 0.0001 | 0.0043 | 0.0007 | 0 | 0 | 0 | 0.0012 | 0.0001 | 0.0009 | 0.0009 | | |
| п | 0.002 | 0 | 0 | 0 | 0 | 0.0015 | 0 | 0 | 0.0002 | 0 | 0 | 0.0004 | 0 | 0.0001 | 0.0085 | 0 | 0.011 | 0 | 0 | 0.0002 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.0001 | 0 | 0 | 0 | 0.0001 | 0 | |
| р | 0.0064 | 0 | 0.0012 | 0 | 0.0055 | 0.0009 | 0 | 0 | 0.0002 | 0.0002 | 0.0003 | 0.0001 | 0.0003 | 0.0003 | 0.0058 | 0.0001 | 0.0001 | 0.0002 | 0.0003 | 0.0001 | 0.0002 | 0.0001 | 0.0001 | 0.0001 | 0.0001 | 0 | 0 | 0.0003 | 0.0001 | 0 | 0 | 0.0001 | 0 | |
| с | 0.0019 | 0 | 0.001 | 0 | 0.0003 | 0.0032 | 0 | 0 | 0.000001 | 0 | 0.000001 | 0.000001 | 0.000001 | 0.000001 | 0.000001 | 0.000001 | 0.000001 | 0.000001 | 0.000001 | 0.000001 | 0.000001 | 0.000001 | 0.000001 | 0.000001 | 0.000001 | 0.000001 | 0.000001 | 0.000001 | 0.000001 | 0.000001 | 0.000001 | 0.000001 | 0.000001 | |
| т | 0.0069 | 0.0005 | 0.0007 | 0.0001 | 0.0011 | 0.0058 | 0.0001 | 0.0001 | 0.0007 | 0.0007 | 0.0008 | 0.0006 | 0.0003 | 0.0027 | 0.0238 | 0.0009 | 0.0021 | 0.0045 | 0.0006 | 0.0004 | 0 | 0 | 0.0008 | 0 | 0 | 0 | 0 | 0.0005 | 0.0091 | 0.0003 | 0.0006 | 0.0006 | | |
| у | 0.0002 | 0.0004 | 0.0009 | 0.0003 | 0.0021 | 0.0011 | 0.0009 | 0.0001 | 0.0007 | 0.0001 | 0.0012 | 0.0001 | 0.0006 | 0.0009 | 0.0005 | 0.0006 | 0.0003 | 0.0009 | 0.0026 | 0 | 0 | 0.0002 | 0.0022 | 0.0002 | 0.0046 | 0 | 0 | 0.0001 | 0.0001 | 0.0013 | 0 | 0 | 0 | |
| ф | 0.0001 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.0003 | 0 | 0 | 0 | 0 | 0.0004 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | |
| х | 0.0002 | 0.0001 | 0.0006 | 0.0001 | 0.0002 | 0.0002 | 0.0001 | 0.0001 | 0.0005 | 0 | 0.0002 | 0.0001 | 0.0002 | 0.0005 | 0.0002 | 0.0005 | 0.0002 | 0.0007 | 0.0003 | 0.0001 | 0.0001 | 0 | 0.0002 | 0 | 0 | 0 | 0 | 0 | 0.0001 | 0 | 0 | 0 | 0 | |
| ц | 0.0005 | 0.0001 | 0 | 0 | 0 | 0.0006 | 0 | 0 | 0.0006 | 0 | 0 | 0.0006 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.0002 | 0 | 0 | 0 | 0 | 0 | |
| ч | 0.0004 | 0 | 0 | 0 | 0.0056 | 0 | 0 | 0.0038 | 0 | 0.0001 | 0.0001 | 0 | 0.0013 | 0 | 0 | 0 | 0 | 0.0071 | 0.0006 | 0 | 0 | 0 | 0 | 0.0001 | 0 | 0 | 0 | 0 | 0.0001 | 0 | 0 | 0 | 0 | |
| ш | 0.0004 | 0 | 0 | 0 | 0.0022 | 0 | 0 | 0.0004 | 0 | 0.0001 | 0 | 0 | 0.0001 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.0003 | 0 | 0 | 0 | 0 | |
| щ | 0.0003 | 0 | 0 | 0 | 0.0052 | 0 | 0 | 0.0019 | 0 | 0 | 0 | 0 | 0 | 0.0015 | 0 | 0 | 0 | 0 | 0 | 0.0001 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.0002 | 0 | 0 | 0 | 0 | |
| ъ | 0 | 0 | 0 | 0 | 0.0001 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.0004 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.0001 | 0.0001 | |
| ы | 0.0001 | 0.0003 | 0.0019 | 0.0002 | 0.0003 | 0.0014 | 0 | 0.0002 | 0.0005 | 0.0005 | 0.0002 | 0.0012 | 0.0021 | 0.0005 | 0.0004 | 0.0005 | 0.0002 | 0.0015 | 0.0002 | 0.0001 | 0.0014 | 0.0002 | 0.0002 | 0.0001 | 0 | 0.0014 | 0 | 0.0002 | 0.0002 | 0 | 0 | 0.0001 | 0.0001 | |
| ь | 0.0002 | 0.0004 | 0.0009 | 0 | 0.0005 | 0.0006 | 0.0001 | 0.0004 | 0.0014 | 0 | 0.0016 | 0.0002 | 0.0004 | 0.0038 | 0.0009 | 0.0008 | 0.0001 | 0.0014 | 0.0007 | 0.0001 | 0.0001 | 0 | 0.0007 | 0.0008 | 0 | 0 | 0 | 0 | 0.0001 | 0.0004 | 0.0001 | 0.0001 | 0.0001 | |
| э | 0 | 0 | 0 | 0 | 0 | 0.0003 | 0 | 0 | 0.0003 | 0 | 0.0003 | 0 | 0.0004 | 0 | 0 | 0 | 0.0038 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.0002 | 0.0002 | 0.0002 | 0.0002 | 0.0002 |
| ю | 0.0001 | 0.0001 | 0.0002 | 0.0002 | 0 | 0 | 0 | 0.0002 | 0 | 0.0001 | 0 | 0.0001 | 0.0001 | 0.0001 | 0.0002 | 0 | 0.0001 | 0.0021 | 0.0001 | 0.0001 | 0 | 0 | 0.0002 | 0.0012 | 0.0012 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| я | 0.0003 | 0.0003 | 0.0017 | 0.0001 | 0.0007 | 0.0007 | 0.0007 | 0.0001 | 0.0003 | 0.0013 | 0.0012 | 0.0003 | 0.0009 | 0.0015 | 0.0008 | 0.0008 | 0.0002 | 0.0013 | 0.0024 | 0.0002 | 0.0005 | 0.0005 | 0.0005 | 0.0008 | 0 | 0 | 0 | 0 | 0.0002 | 0.0002 | 0.0002 | 0.0002 | 0.0002 | |

Ентропія та надлишковість

N1, N2

```
З пробілами:  
N1 (монограми): 4.2996, R1: 0.1549  
N2 (біграми, перекриття): 3.8100, R2: 0.2511  
N2 (біграми, без перекриття): 3.8091, R2: 0.2513  
  
Без пробілів:  
N1 (монограми): 4.3746, R1: 0.1328  
N2 (біграми, перекриття): 3.9773, R2: 0.2115  
N2 (біграми, без перекриття): 3.9766, R2: 0.2117
```

N10

Лабораторная работа №1



Произвольная часть текста:
то_что_я_

Использованные буквы:

Порядок n-граммы:
5 символов
10 символов
15 символов
20 символов
25 символов
30 символов
35 символов
40 символов
45 символов
50 символов

Введенный символ:
Символ по счету:
Номер эксперимента: 51
Поле ввода символов:

Продолжить

Другой

Неравенство для энтропии:
3.06252289332512 < H < 3.75457162242399

Двоичная таблица угаданных символов:
00000001000000000000000000000000
10000000000000000000000000000000
10000000000000000000000000000000
00000000000000000000000001000000
00000000000000000000000000000000
.....

Вероятности:
q[1] = 0.28
q[2] = 0.14
q[3] = 0.06
q[4] = 0.06
q[5] = 0
q[6] = 0.04
q[7] = 0.04
q[8] = 0.02
q[9] = 0.02
q[10] = 0
q[11] = 0.02
q[12] = 0
q[13] = 0.02
q[14] = 0.04
q[15] = 0.02
q[16] = 0.02
q[17] = 0.02
q[18] = 0.04
q[19] = 0
q[20] = 0.02
q[21] = 0.04
q[22] = 0
q[23] = 0
q[24] = 0
q[25] = 0
q[26] = 0.04
q[27] = 0.02
q[28] = 0
q[29] = 0
q[30] = 0
q[31] = 0.02
q[32] = 0.02

Строка состояния:

0.3251 < R10 < 0.3875

Н20

Лабораторная работа №1



Произвольная часть текста:
целый_рой_извинений

Использованные буквы:

Порядок n-граммы:
5 символов
10 символов
15 символов
20 символов
25 символов
30 символов
35 символов
40 символов
45 символов
50 символов

Введенный символ:
Символ по счету:
Номер эксперимента: 51
Поле ввода символов:

Продолжить Другой

Неравенство для энтропии:
2.16854562214316 < H < 2.74450175314363
Двоичная таблица угаданных символов:
00000000001000000000000000000000
00000000001000000000000000000000
00000000000000000000000000000001
10000000000000000000000000000000
01000000000000000000000000000000

Вероятности:
q[1] = 0.42
q[2] = 0.12
q[3] = 0.2
q[4] = 0
q[5] = 0
q[6] = 0.02
q[7] = 0
q[8] = 0
q[9] = 0.04
q[10] = 0
q[11] = 0.04
q[12] = 0
q[13] = 0.02
q[14] = 0.02
q[15] = 0
q[16] = 0.02
q[17] = 0
q[18] = 0
q[19] = 0.02
q[20] = 0
q[21] = 0.02
q[22] = 0
q[23] = 0
q[24] = 0
q[25] = 0
q[26] = 0
q[27] = 0
q[28] = 0
q[29] = 0
q[30] = 0.02
q[31] = 0.02
q[32] = 0.02

Строка состояния:

$$0.4511 < R20 < 0.5663$$

Н30

Лабораторная работа №1



Произвольная часть текста:
к_следующему_он_состоит_в_том_что_никто_из_нас_по_настоящему_не_следует_зак

Использованные буквы:

Порядок n-граммы:
5 символов
10 символов
15 символов
20 символов
25 символов
30 символов
35 символов
40 символов
45 символов
50 символов

Введенный символ: _ (пробел)
Символ по счету: 1
Номер эксперимента: 50
Поле ввода символов:

Продолжить Другой

Неравенство для энтропии:
 $1.40380437162245 < H < 2.07126202070915$
Двоичная таблица угаданных символов:
00010000000000000000000000000000
00100000000000000000000000000000
01000000000000000000000000000000
10000000000000000000000000000000
00010000000000000000000000000000
00000000000000000000000000000000

Вероятности:
q[1] = 0.52
q[2] = 0.22
q[3] = 0.06
q[4] = 0.1
q[5] = 0.04
q[6] = 0
q[7] = 0.02
q[8] = 0
q[9] = 0
q[10] = 0
q[11] = 0
q[12] = 0
q[13] = 0
q[14] = 0
q[15] = 0
q[16] = 0
q[17] = 0
q[18] = 0
q[19] = 0
q[20] = 0
q[21] = 0
q[22] = 0
q[23] = 0
q[24] = 0
q[25] = 0
q[26] = 0.02
q[27] = 0
q[28] = 0.02
q[29] = 0
q[30] = 0
q[31] = 0
q[32] = 0

Строка состояния:
Вы угадали. Для продолжения опыта нажмите "Продолжить", или "Другой" для выбора другого порядка

0.5857 < R30 < 0.7192

Висновок

В результаті виконання комп'ютерного практикуму було вивчено поняття ентропії джерела тексту, частотного розподілу символів та біграм, їх надлишковості. Було створено програму для підрахунку частот літер та біграм у тексті, обчислення ентропії, виведення частот у вигляді таблиць та матриць, обчислення надлишковості.

Використання пробілів знижує ентропію оскільки пробіли є одними найрозповсюдженіших символів. Ентропія біграм також зменшується у моделі без перекриття через меншу кількість унікальних пар символів. Це вказує на високий рівень надлишковості російської мови, що є характерним для природних мов.

За допомогою програми CoolPinkProgram було проведено експеримент з визначенням умовної ентропії для $H(10)$, $H(20)$, $H(30)$. Експеримент підтвердив основну властивість природних мов: збільшення довжини n -грам зменшує умовну ентропію $H(n)$ завдяки кращому врахуванню закономірностей тексту. Надлишковість R , у свою чергу, зростає, що показує, наскільки структурованим і передбачуваним є текст.