Міністерство освіти і науки, молоді та спорту України
Національний технічний університет України
“Київський політехнічний інститут”
Фізико-Технічний інститут

Лабораторна робота №1
з Симетричної криптографії

Студент групи ФІ-93
Карловський Володимир Олександрович

Задачі

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.

№1

Для отримання детального результату треба запустити програму (інструкція в README.md) не розміщую результат тут бо в російському алфавіті 33 літери => 33*33 = 1089 біграм

Короткі результати ентропій:

```
statByLettersWithSpaces has entropy: 4.4953369017542135
statByLettersWithoutSpaces has entropy: 4.462151364016589
statByBigramsWithIntersAndWithoutSpaces has entropy: 4.1453958533960975
statByBigtamsWithIntersAndWithSpaces has entropy: 3.979123994939124
statByBigramsWithoutIntersAndWithoutSpaces has entropy: 4.145212314083562
statByBigtamsWithoutIntersAndWithSpaces has entropy: 3.937872933272794
```

Тобто

$H_{1.1} = 4,49$ – ентропія літер з пробілами

$H_{1.2} = 4,46$ – ентропія без пробілів

$H_{2.1} = 4,14$ – ентропія біграм з перетином і без пробілів

$H_{2.2} = 3,97$ – ентропія біграм з перетином і з пробілами

$H_{2.3} = 4,14$ – ентропія біграм без перетину і без пробілів

$H_{2.4} = 3,93$ – ентропія біграм без перетину і з пробілами

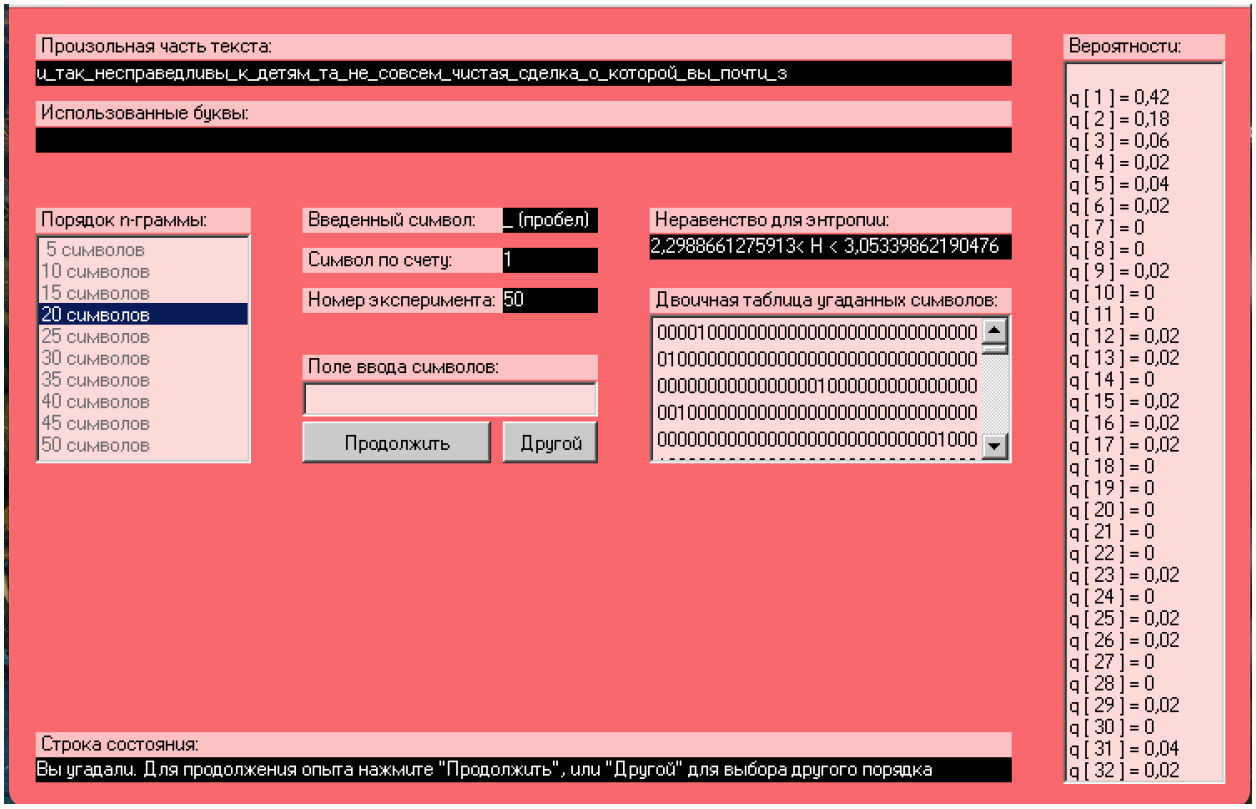
№2

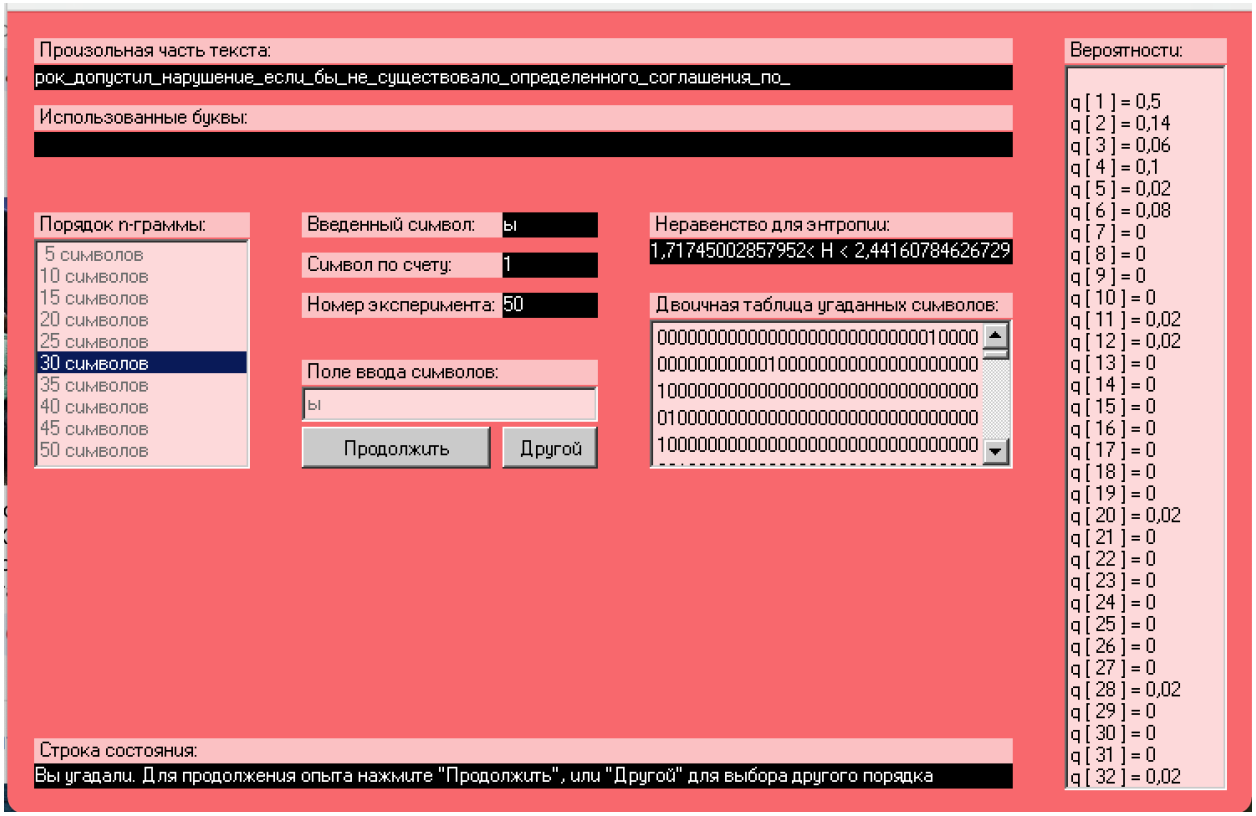
Результати:

$2,23 < H^{(10)} < 2,96$

$2,29 < H^{(20)} < 3,05$

$1,71 < H^{(30)} < 2,44$


Произольная часть текста:
и_так_несправедливы_к_детям_та_не_совсем_чистая_сделка_о_которой_вы_почти_з
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:
2,2988661275913< H < 3,05339862190476
Двоичная таблица угаданных символов:
Строка состояния:
Вы угадали. Для продолжения опыта нажмите "Продолжить", или "Другой" для выбора другого порядка
Вероятности:
q [1] = 0,42
q [2] = 0,18
q [3] = 0,06
q [4] = 0,02
q [5] = 0,04
q [6] = 0,02
q [7] = 0
q [8] = 0
q [9] = 0,02
q [10] = 0
q [11] = 0
q [12] = 0,02
q [13] = 0,02
q [14] = 0
q [15] = 0,02
q [16] = 0,02
q [17] = 0,02
q [18] = 0
q [19] = 0
q [20] = 0
q [21] = 0
q [22] = 0
q [23] = 0,02
q [24] = 0
q [25] = 0,02
q [26] = 0,02
q [27] = 0
q [28] = 0
q [29] = 0,02
q [30] = 0
q [31] = 0,04
q [32] = 0,02


Произольная часть текста:
рок_допустил_нарушение_если_бы_не_существовало_определенного_соглашения_по_
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: ы
Символ по счету: 1
Номер эксперимента: 50
Поле ввода символов:
ы
Продолжить
Другой
Неравенство для энтропии:
1,71745002857952< H < 2,44160784626729
Двоичная таблица угаданных символов:
Строка состояния:
Вы угадали. Для продолжения опыта нажмите "Продолжить", или "Другой" для выбора другого порядка
Вероятности:
q [1] = 0,5
q [2] = 0,14
q [3] = 0,06
q [4] = 0,1
q [5] = 0,02
q [6] = 0,08
q [7] = 0
q [8] = 0
q [9] = 0
q [10] = 0
q [11] = 0,02
q [12] = 0,02
q [13] = 0
q [14] = 0
q [15] = 0
q [16] = 0
q [17] = 0
q [18] = 0
q [19] = 0
q [20] = 0,02
q [21] = 0
q [22] = 0
q [23] = 0
q [24] = 0
q [25] = 0
q [26] = 0
q [27] = 0
q [28] = 0,02
q [29] = 0
q [30] = 0
q [31] = 0
q [32] = 0,02

№3

$H_0 = log_2 32 = 5$

$H_{inf} \approx H^{(30)}$

$0.66 > R > 0.51$

$H_{inf} \approx H^{(20)}$

$0.54 > R > 0.39$

$H_{inf} \approx H^{(10)}$

$0.55 > R > 0.40$

$H_{inf} \approx H_{2.4}$

$R = 0.22$

$H_{inf} \approx H_{2.3}$

$R = 0.17$

$H_{inf} \approx H_{2.2}$

$R = 0.20$

$H_{inf} \approx H_{2.1}$

$R = 0.17$

$H_{inf} \approx H_1$

$R = 0.10$