

НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ УКРАЇНИ «КИЇВСЬКИЙ ПОЛІТЕХНІЧНИЙ ІНСТИТУТ ІМЕНІ ІГОРЯ СІКОРСЬКОГО» ФІЗИКО-ТЕХНІЧНИЙ ІНСТИТУТ

Кафедра Інформаційної Безпеки

Лабораторна робота №1 дисципліни

"КРИПТОГРАФІЯ"

Підготував:

студент групи ФБ-06

Жак Костянтин

Тема роботи: Експериментальна оцінка ентропії на символ джерела відкритого тексту

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

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 \square_I та \square_2 за безпосереднім означенням. Підрахувати частоти букв та біграм		
а також значення \square_I та \square_2 на довільно обраному тексті російською мовою достатньої		
довжини (щонайменше 1Мб), де імовірності замінити відповідними частотами. Також		
одержати значення \Box_I та \Box_2 на тому ж тексті, в якому вилучено всі пробіли.		
2. За допомогою програми CoolPinkProgram оцінити значення $\Box^{(10)}$, $\Box^{(20)}$, $\Box^{(30)}$.		
3. Використовуючи отримані значення ентропії, оцінити надлишковість російської		
мови в різних моделях джерела.		

Хід роботи

За початковий текст взяв "Хроніки нарнії" Значення частот букв з пробілами:

Буква	Ïï частотність
a	0.06865
6	0.01302
В	0.03577
г	0.0144
д	0.0275
е	0.0623
ж	0.00803
3	0.01467
И	0.05746
й	0.00799
К	0.02897
л	0.04541
М	0.02643
н	0.05237
0	0.09153
п	0.02322
р	0.03641
С	0.04534
т	0.04996
у	0.02378
ф	0.00066
X	0.00749
ц	0.00244
Ч	0.01247
ш	0.00784
щ	0.00225
ы	0.01605
ь	0.01802
3	0.00314
Ю	0.0049
Я	0.0159
пробел	0.17026

4.354901763494896

Надлишковість:

0.1290196473010209

Значення частот букв без пробілів:

Буква а 0.08274 6 0.01569 в 0.04312 г 0.01735 д 0.03315 е 0.07508 ж 0.00968 з 0.01768 и 0.06925 й 0.00963 к 0.03491 л 0.05473 м 0.03185 н 0.06312 ο 0.11031 п 0.02799 р 0.04388 c 0.05465 т 0.06021 у 0.02866 ф 0.0008 х 0.00903 ц 0.00294 ч 0.01503 ш 0.00294 ч 0.01503 ш 0.00271 ы 0.00271 ы 0.00271 ы 0.00271 ы 0.00272 э 0.00378 ю 0.0059 я 0.0059		
6 0.01569 В 0.04312 Г 0.01735 Д 0.03315 е 0.07508 ж 0.00968 З 0.01768 И 0.06925 Й 0.00963 К 0.03491 Л 0.05473 М 0.03185 Н 0.06312 0 0.11031 П 0.02799 р 0.04388 c 0.05465 Т 0.06021 У 0.02866 Ф 0.0088 х 0.09903 Ц 0.00294 Ч 0.01503 Ш 0.00271 Ы 0.00271 Ы 0.00271 Ы 0.00271 Ы 0.00271 Ы 0.001503	Буква	
В 0.04312 Г 0.01735 Д 0.03315 е 0.07508 ж 0.00968 з 0.01768 и 0.06925 й 0.00963 к 0.03491 Л 0.05473 м 0.03185 H 0.06312 0 0.11031 П 0.02799 р 0.04388 c 0.05465 T 0.06021 y 0.02866 ф 0.0098 x 0.00903 Ц 0.00294 Ч 0.01503 Ш 0.00271 Ы 0.00271 Ы 0.00935 Ь 0.001722 9 0.00378 Ю 0.0059	a	0.08274
Г	6	0.01569
Д 0.03315 e 0.07508 ж 0.00968 з 0.01768 и 0.06925 й 0.00963 к 0.03491 л 0.05473 м 0.03185 н 0.06312 о 0.11031 п 0.02799 р 0.04388 с 0.05465 т 0.06021 у 0.02866 ф 0.0098 х 0.00903 ц 0.00294 ч 0.01503 ш 0.00271 ы 0.00271 ы 0.00271 ы 0.00271 ы 0.00271 ы 0.00378 ю 0.0059	В	0.04312
е 0.07508 ж 0.00968 з 0.01768 и 0.06925 й 0.00963 к 0.03491 л 0.05473 м 0.03185 н 0.06312 о 0.11031 п 0.02799 р 0.04388 с 0.05465 т 0.06021 у 0.02866 ф 0.0098 х 0.00903 ц 0.00294 ч 0.01503 ш 0.00294 ч 0.01503 ш 0.00271 ы 0.00378 ю 0.0059	Γ	0.01735
ж 0.00968 3 0.01768 и 0.06925 й 0.00963 к 0.03491 л 0.05473 м 0.03185 н 0.06312 о 0.11031 п 0.02799 р 0.04388 с 0.05465 т 0.06021 у 0.02866 ф 0.0008 х 0.00903 ц 0.00993 ц 0.00294 ч 0.01503 ш 0.00944 щ 0.00271 ы 0.00971 ы 0.01935 ь 0.02172 э 0.00378 ю 0.0059	Д	0.03315
3	е	0.07508
и 0.06925 й 0.00963 к 0.03491 л 0.05473 м 0.03185 н 0.06312 о 0.11031 п 0.02799 р 0.04388 с 0.05465 т 0.06021 у 0.02866 ф 0.00903 ц 0.00294 ч 0.00294 щ 0.00944 щ 0.01503 ы 0.02172 э 0.00378 ю 0.0059	ж	0.00968
й 0.00963 к 0.03491 л 0.05473 м 0.03185 н 0.06312 о 0.11031 п 0.02799 р 0.04388 с 0.05465 т 0.06021 у 0.02866 ф 0.0988 х 0.00903 ц 0.01503 ш 0.00294 ц 0.00944 ц 0.00271 ы 0.02172 э 0.00378 ю 0.0059	3	0.01768
К 0.03491 Л 0.05473 М 0.03185 Н 0.06312 О 0.11031 П 0.02799 Р 0.04388 С 0.05465 Т 0.06021 У 0.02866 Ф 0.0098 X 0.00903 Ц 0.00294 Ч 0.01503 Ш 0.00271 Ы 0.00271 Ы 0.01935 Ь 0.02172 Э 0.00378 Ю 0.0059	И	0.06925
л 0.05473 м 0.03185 н 0.06312 о 0.11031 п 0.02799 р 0.04388 с 0.05465 т 0.06021 у 0.02866 ф 0.0008 х 0.00903 ц 0.00294 ч 0.01503 ш 0.00944 щ 0.00271 ы 0.00271 ы 0.01935 ь 0.02172 э 0.00378	Й	0.00963
м 0.03185 н 0.06312 о 0.11031 п 0.02799 р 0.04388 с 0.05465 т 0.06021 у 0.02866 ф 0.0098 х 0.00903 ц 0.00294 ч 0.01503 ш 0.00271 ы 0.00271 ы 0.01935 ь 0.02172 э 0.00378 ю 0.0059	К	0.03491
H 0.06312 0 0.11031 П 0.02799 р 0.04388 с 0.05465 Т 0.06021 у 0.02866 ф 0.0008 х 0.00903 Ц 0.00294 Ч 0.01503 Ш 0.00944 Щ 0.00271 Ы 0.01935 Ь 0.02172 Э 0.00378 Ю 0.0059	Л	0.05473
0 0.11031 п 0.02799 р 0.04388 с 0.05465 т 0.06021 у 0.02866 ф 0.0098 х 0.00993 ц 0.01503 ш 0.005944 щ 0.01935 ь 0.02172 э 0.00378 ю 0.0059	М	0.03185
П 0.02799 р 0.04388 с 0.05465 Т 0.06021 у 0.02866 ф 0.0008 х 0.00903 ц 0.00294 ч 0.01503 ш 0.00944 щ 0.00271 ы 0.01935 ь 0.02172 э 0.00378 ю 0.0059	Н	0.06312
р 0.04388 c 0.05465 т 0.06021 y 0.02866 ф 0.0008 x 0.00903 U 0.00294 U 0.01503 U 0.00944 U 0.00271 Ы 0.01935 Ь 0.02172 Э 0.00378 Ю 0.0059	0	0.11031
С 0.05465 Т 0.06021 У 0.02866 Ф 0.0008 X 0.00903 Ц 0.00294 Ч 0.01503 Ш 0.00944 Щ 0.00271 Ы 0.01935 Ь 0.02172 Э 0.00378 Ю 0.0059	П	0.02799
Т 0.06021 У 0.02866 Ф 0.0008 X 0.00903 Ц 0.00294 Ч 0.01503 Ш 0.00944 Щ 0.00271 Ы 0.01935 Ь 0.02172 Э 0.00378 Ю 0.0059	p	0.04388
у 0.02866 ф 0.0008 x 0.00903 Ц 0.00294 Ч 0.01503 Ш 0.00944 Щ 0.00271 Ы 0.01935 Ь 0.02172 Э 0.00378	С	0.05465
ф 0.0008 x 0.00903 ц 0.00294 ч 0.01503 ш 0.00944 щ 0.00271 ы 0.01935 ь 0.02172 э 0.00378 ю 0.0059	T	0.06021
X 0.00903 Ц 0.00294 Ч 0.01503 Ш 0.00944 Щ 0.00271 Ы 0.01935 Ь 0.02172 Э 0.00378 Ю 0.0059	у	0.02866
Ц 0.00294 Ч 0.01503 Ш 0.00944 Щ 0.00271 Ы 0.01935 Ь 0.02172 Э 0.00378 Ю 0.0059	ф	0.0008
Ч 0.01503 Ш 0.00944 Щ 0.00271 Ы 0.01935 Ь 0.02172 Э 0.00378 Ю 0.0059	X	0.00903
ш 0.00944 щ 0.00271 ы 0.01935 ь 0.02172 э 0.00378	ц	0.00294
Щ 0.00271 ы 0.01935 ь 0.02172 э 0.00378 ю 0.0059	Ч	0.01503
ы 0.01935 ь 0.02172 э 0.00378 ю 0.0059	Ш	0.00944
ь 0.02172 э 0.00378 ю 0.0059	Щ	0.00271
э 0.00378 ю 0.0059	Ы	0.01935
ю 0.0059	Ь	0.02172
	Э	0.00378
я 0.01917	Ю	0.0059
	Я	0.01917

4.45752152505319

Надлишковість:

0.10025335174796479

Значення частот біграм що перетинаються з пробілами:

3.968282216187312

Надлишковість:

0.20634355676253757

Значення частот біграм що перетинаються без пробілів:

Ентропія:

3.968282216187312

Надлишковість:

0.19900585936260018

Значення частот біграм що не переитнаються з пробілами:

3.9689715592964325

Надлишковість:

0.2062056881407135

Значення частот біграм що не перетинаються без пробілів:

Ентропія:

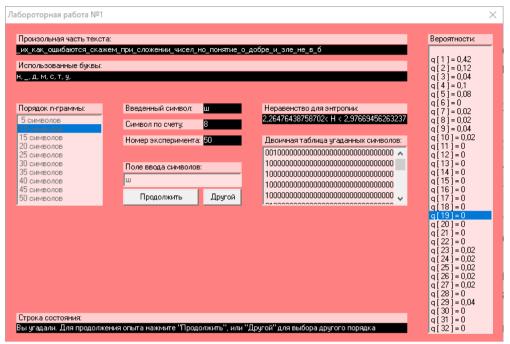
3.9689715592964325

Надлишковість:

0.19886671608568252

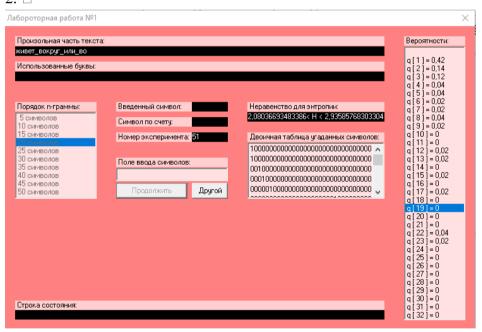
2. CoolPinkProgram

1. □⁽¹⁰⁾



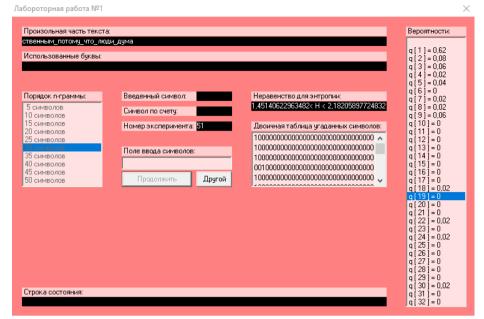
 $2.265 < \Box^{(10)} < 2.980$

2. □⁽²⁰⁾



 $2.080 < \Box^{(20)} < 2.940$

3. □⁽³⁰⁾



 $1.451 < \Box^{(30)} < 2.182$

3. Оцінимо надлишковість на основі отриманих результатів:

Для нижньої границі:

1)
$$R1 = 1 - \frac{2.26476438758702}{5} = 0.547047122482596$$

- 2) R2 = 0.583926613033228
- 3) R3 = 0.709718754073036

Для верхньої границі:

- 1) R1 = 0.404661087473526
- 2) R2 = 0.412828463393392
- 3) R3 = 0.563588204550336

Висновок

Під час виконання даного лабораторного практикуму було написано програму, що обраховує частотність літер та біграм російської мови на основі тексту "Хроніки Нарнії", на основі цих результатів було обраховано надлишковість російської мови. У шифруванні слід використовувати текст без пробілів, так як ентропія в ньому більша.

Також навчився працювати з програмою CoolPinkProgram.