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:

Акент'єв Влад, Шапоренко Микита

Група: ФБ-06

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.

Хід роботи

Частота літер з пробілами

	0	1
0		0,115798
1	о	0,067325
2	е	0,049301
3	а	0,048312
4	н	0,039437
5	и	0,038972
6	т	0,036785
7	л	0,032214
8	с	0,031571
9	р	0,028385
10	в	0,025004
11	м	0,021576
12	к	0,020979
13	д	0,017241
14	у	0,016918
15	п	0,01617
16	я	0,015931
17	ы	0,012972
18	г	0,012265
19	ь	0,011822
20	з	0,010751
21	б	0,010365
22	ч	0,008443
23	й	0,007134
24	ж	0,006186
25	х	0,005646
26	ш	0,005085
27	ю	0,003491
28	ц	0,002401
29	щ	0,001818
30	э	0,00171
31	ф	0,000564
32	ъ	0,000098

Частота літер без пробілів

	0	1
0	о	0,110938
1	е	0,081238
2	а	0,079608
3	н	0,064984
4	и	0,064218
5	т	0,060614
6	л	0,053082
7	с	0,052023
8	р	0,046773
9	в	0,041201
10	м	0,035553
11	к	0,034569
12	д	0,02841
13	у	0,027877
14	п	0,026645
15	я	0,026251
16	ы	0,021375
17	г	0,02021
18	ь	0,01948
19	з	0,017715
20	б	0,017079
21	ч	0,013912
22	й	0,011755
23	ж	0,010193
24	х	0,009303
25	ш	0,008379
26	ю	0,005752
27	ц	0,003956
28	щ	0,002996
29	э	0,002818
30	ф	0,000929
31	ъ	0,000161

```
H1 для літер з пробілами:  4.390237487048099
H1 для літер без пробілів:  4.471928261102778
Надлишковість для H1:  0.12967992128131944
Надлишковість для H1 без пробілів:  0.10561434777944445
```

Частота біграм з пробілами

	0	1
0	о	0,013507
1	и	0,012574
2	с	0,011883
3	е	0,011647
4	а	0,011244
5	п	0,010961
6	н	0,01065
7	я	0,010638
8	в	0,009773
9	то	0,00811
10	ст	0,007532
11	на	0,007313
12	ь	0,007253
13	м	0,007176
14	о	0,007123
15	но	0,00662
16	к	0,006588
17	и	0,006516
18	ал	0,006416
19	по	0,006395
20	не	0,006363
21	л	0,006083
22	го	0,005869
23	ра	0,005839
24	ко	0,005695
25	ро	0,005391
26	ка	0,005296
27	т	0,005259
28	ен	0,005255
29	й	0,005251
30	от	0,005172

Частота біграм з пробілами кроком 2

	0	1
0	о	0,006647
1	и	0,006162
2	с	0,005962
3	е	0,005789
4	а	0,005668
5	п	0,0055
6	я	0,005379
7	н	0,005359
8	в	0,004915
9	то	0,004077
10	ст	0,003749
11	м	0,003688
12	на	0,003669
13	ь	0,003633
14	о	0,003601
15	к	0,003368
16	но	0,003343
17	и	0,00328
18	ал	0,003249
19	по	0,003224
20	не	0,003173
21	ра	0,002988
22	л	0,002971
23	го	0,002969
24	ко	0,002821
25	й	0,002677
26	ен	0,002663
27	ро	0,002645
28	от	0,002626
29	т	0,002575
30	ка	0,002544
31	м	0,002534

```
H2 для біграм з пробілами:  3.9972273941022305
H2 для біграм з пробілами кроком 2:  3.996669608586941
Надлишковість для H2 з пробілами:  0.20759018833155762
Надлишковість для H2 з пробілами кроком 2:  0.20770076365578705
```

Частота біграм без пробілів

	0	1
0	то	0,01675
1	ст	0,01541
2	на	0,01468
3	но	0,01353
4	ал	0,0132
5	по	0,0128
6	не	0,01279
7	ен	0,01241
8	го	0,01184
9	ко	0,01177
10	ра	0,01169
11	от	0,01169
12	ос	0,01122
13	ов	0,01115
14	ро	0,01084
15	ка	0,01063
16	он	0,01038
17	ли	0,01011
18	ло	0,00998
19	ор	0,00995
20	ни	0,00992
21	ол	0,00952
22	ть	0,00948
23	ер	0,00938
24	во	0,00931
25	ре	0,00889
26	пр	0,00881
27	ел	0,0088
28	ат	0,00862
29	ес	0,00852
30	ин	0,00851
31	ом	0,00843

Частота біграм без пробілів кроком 2

	0	1
0	то	0,00831
1	ст	0,0077
2	на	0,007252
3	но	0,006834
4	ал	0,006476
5	ен	0,00638
6	не	0,006334
7	по	0,006314
8	го	0,006024
9	ко	0,005888
10	от	0,005856
11	ра	0,0058
12	ос	0,005692
13	ов	0,00563
14	ро	0,005394
15	ка	0,005196
16	ли	0,005112
17	он	0,00511
18	ло	0,00501
19	ни	0,005
20	ор	0,004918
21	ол	0,00485
22	во	0,004732
23	ер	0,00465
24	ть	0,004592
25	ел	0,004456
26	ат	0,004418
27	ре	0,004412
28	пр	0,004282
29	ин	0,004242
30	ла	0,004178
31	ес	0,004156

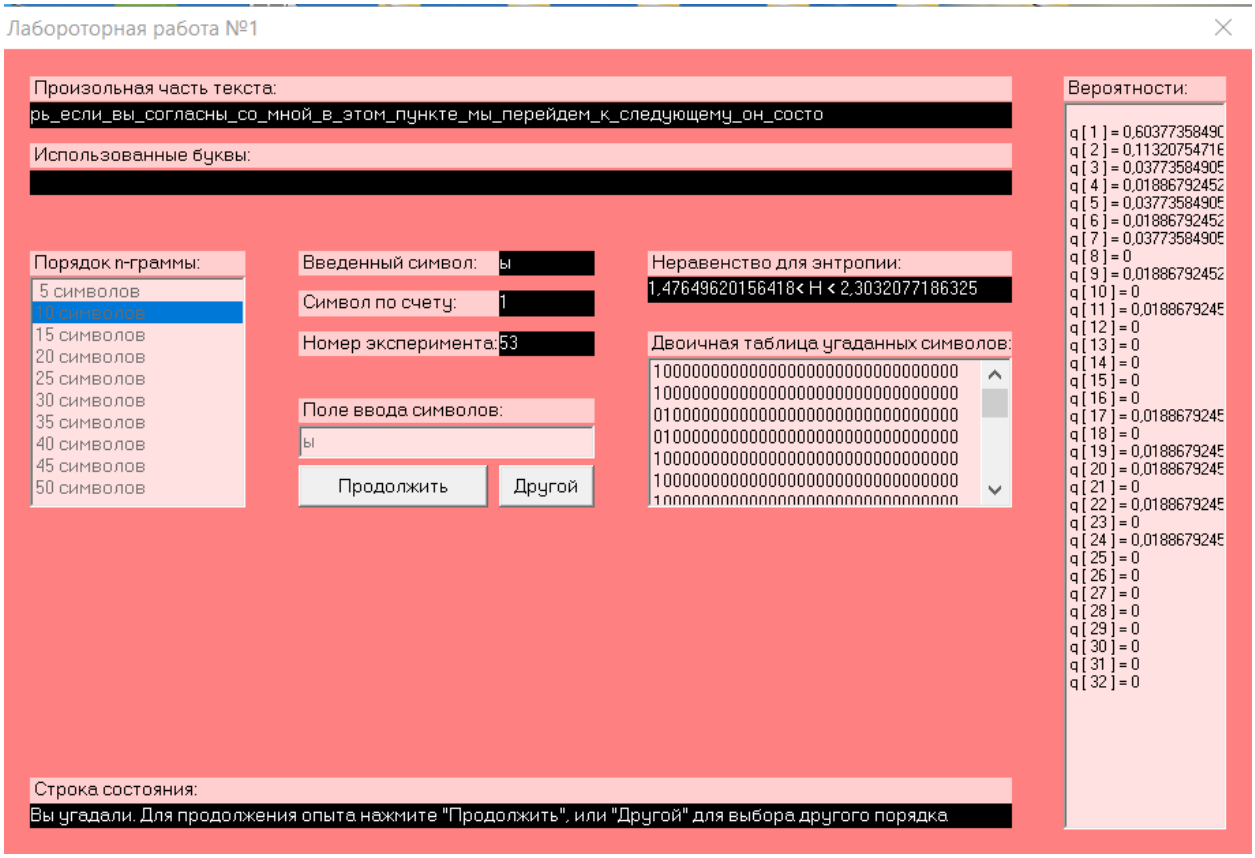
```
H2 для біграм без пробілів:  4.163743197168655
H2 для біграм без пробілів кроком 2:  4.163474531693412
Надлишковість для H2 без пробілів:  0.167251360566269
Надлишковість для H2 без пробілів кроком 2:  0.16730509366131763
```

Таблиця значень

	H1		H2 з пробілами		H2 без пробілів	
	з пробілами	без пробілів	кроком 1	кроком 2	кроком 1	кроком 2
H	4.39023	4.47192	3.99722	3.99666	4.16374	4.16347
R	0.12967	0.10561	0.20759	0.20770	0.16725	0.16730

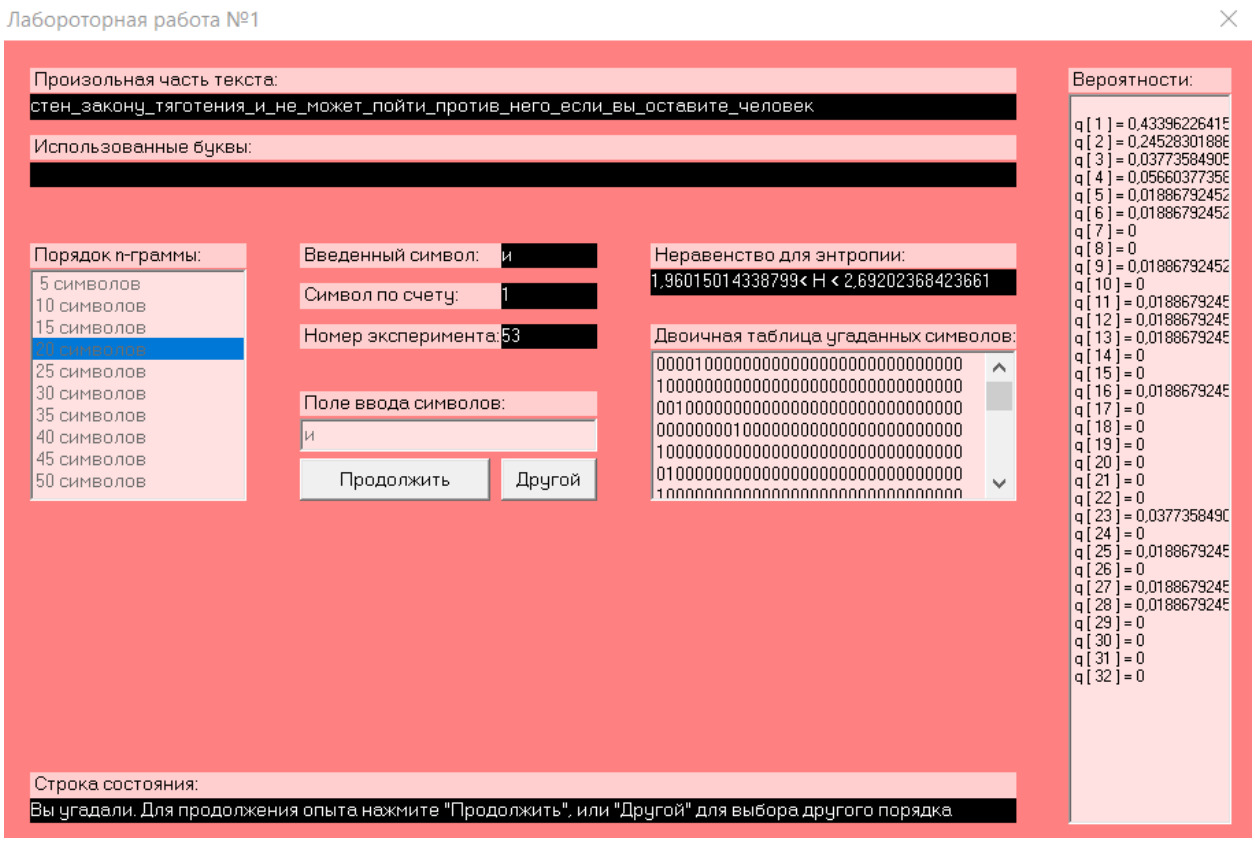
CoolPinkProgram

Для H^10



1.47649620 < H^10 < 2.30320771

Для H^20



1.96015014 < H^20 < 2.69202368

Для H^30

1.81557037 < H^30 < 2.66283881

Таблиця значень

	H^10	H^20	H^30
H	2.30320	2.69202	2.66283
R	0.54341	0.46633	0.47212

Висновки:

В цій лабораторній роботі познайомилися з поняттями ентропії та надлишковості. Навчилися рахувати частоту літер, частоту біграм та оцінювати ентропію. Використовував такі програми в лабораторній роботі: PyCharm та CoolPinkProgram.