



Міністерство освіти і науки України  
Національний технічний університет України  
«Київський політехнічний інститут імені Ігоря Сікорського»  
Фізико-технічний інститут

**ЛАБОРАТОРНА РОБОТА №1**  
**з дисципліни**  
**«Криптографія»**  
**на тему: «Експериментальна оцінка ентропії на символ джерела відкритого тексту»**

Виконав:  
студент 3 курсу ФТІ  
групи ФБ-74  
Сизов Ігор  
Перевірили:  
Чорний О.  
Савчук М. М.  
Завадська Л. О.

## 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  $H_1$  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.

## Результати виконання

	<b>H</b>	<b>H(без пробілів)</b>	<b>R</b>
<b>Монограмми</b>	4,376	4,479	0,16
<b>Біграмми з перетином</b>	4,004	4,127	0,2707
<b>Біграмми без перетину</b>	4,003	4,126	0,27

## Частоти Монограмм

а 0,07

б 0,014

в 0,038

г 0,017

д 0,025

е 0,066

ё 0,001

ж 0,008

з 0,015

и 0,055  
й 0,01  
к 0,03  
л 0,042  
м 0,025  
н 0,054  
о 0,095  
п 0,021  
р 0,038  
с 0,044  
т 0,047  
у 0,024  
ф 0,002  
х 0,007  
ц 0,003  
ч 0,011  
ш 0,008  
щ 0,002  
ъ 0  
ы 0,016  
ь 0,016  
э 0,003  
ю 0,005  
я 0,019  
‘ ‘ 0,165

### Частоти біграмм за перетином

сп 0,00176072	н 0,01448715	кр 0,00179172
па 0,0012108	ик 0,00172482	сн 0,00073594
ас 0,00389022	ко 0,00857839	ым 0,00151595
си 0,00172482	ол 0,00703634	ем 0,00389675
иб 0,00051075	ас 0,0008926	ез 0,00112921
бо 0,00226494	ви 0,00341863	з 0,00106557
ч 0,00513202	ич 0,00121733	зл 0,00018439
чт 0,00356549	ч 7,506E-05	аф 0,0005287

то	0,01385727	лс	0,00140988	ф	0,00011586
о	0,01758757	ст	0,01086944	т	0,00329461
с	0,01577138		4,9E-06	ид	0,00179335
ск	0,00486114	л	0,00268758	уч	0,00072126
ка	0,00784734	ть	0,00510591	чш	7,343E-05
ач	0,00076695	ь	0,00762868	рс	0,0002464
ча	0,00221599	п	0,01545318	ив	0,00270716
ал	0,00925069	пе	0,00253419	ьн	0,00105088
ли	0,00661207	ер	0,00706571	иш	0,00037042
и	0,01498485	рв	0,00026762	шк	0,00042753
к	0,01089229	ва	0,00603767	бу	0,00119774
кн	0,00237427	ая	0,00289971	уд	0,00199896
ни	0,00686989	по	0,00889496	нь	0,00104109
нг	0,00050586	ме	0,00281812	де	0,00446625
гу	0,00081264	ья	0,00039816	ьн	0,00036063
у	0,00412683	мо	0,00366177	нч	0,00019745
в	0,01505012	ве	0,00436507	бя	0,00048465
в	0,00537026	рн	0,00066904	мь	2,448E-05
б	0,00748671	ны	0,00317223	бю	0,00037042
бе	0,00206586	ый	0,00156	ю	0,00239712
ес	0,00358507	р	0,00440261	ок	0,00173787
пл	0,00074247	аб	0,00069352	см	0,00101498
ла	0,00669202	ну	0,00275612	му	0,00227147
ат	0,00485625	ня	0,00348716	тя	0,00051075
тн	0,00091707	яз	0,0015959	яс	0,00090239
но	0,0087122	зь	0,00113574	ою	0,00076695
ой	0,00375478	г	0,00515813	ош	0,00134297
й	0,00579616	ге	0,00058255	дш	0,0001077
э	0,00242649	уа	1,958E-05	ши	0,00200059
эл	8,812E-05	лу	0,00122712	ий	0,00148005
ле	0,00382005	ук	0,00101988	дв	0,00095624
ек	0,00102477	кк	9,79E-06	ун	0,00031004
кт	0,0003949	не	0,00904182	нд	0,00117163
тр	0,00292909	ль	0,00359976	чу	0,00059724
ро	0,00713751	ьш	0,00042264	ул	0,00155511
он	0,00688295	ше	0,00225842	лк	0,00081753
нн	0,00298457	ак	0,00447277	ба	0,00119448
би	0,00077021	к	0,00440587	шм	1,958E-05
бл	0,00079142	ям	0,00033942	св	0,00178193
ио	0,00028557	фа	0,00018603	тл	0,0002823
от	0,00670671	ам	0,00279691	лы	0,00126791
те	0,00475671	ил	0,00511733	ыр	0,00036226
ке	0,00049607	ии	0,00065109	ца	0,00058908
е	0,01284882	ап	0,00075389	ьс	0,00080774
вс	0,00310859	ар	0,00280017	нц	0,00059234

се 0,00279691	рт 0,00066741	цу 0,00050912
ги 0,00085017	ет 0,00367808	уз 0,00089749
а 0,00278549	я 0,00145067	зс 0,00015339
ав 0,00357691	я 0,00873994	ык 0,00018929
вт 0,00020887	с 0,00279854	ум 0,00102477
ор 0,00642278	ре 0,00620085	ды 0,00068536
ра 0,00869751	ед 0,0025195	хи 0,00015992
эт 0,00226821	ду 0,00167423	ел 0,00620412
та 0,00561503	уп 0,0005695	ьс 0,00091381
а 0,01253714	еж 0,00097255	ац 8,975E-05
ж 0,00171829	жд 0,00065272	ци 0,00023335
же 0,00387553	да 0,00420353	ры 0,00149147
га 0,00105578	аю 0,00081101	ые 0,00093176
д 0,00684541	е 0,00387227	йс 0,00068373
др 0,00212624	сл 0,00240038	ус 0,0014523
ру 0,00319018	вы 0,0027708	ку 0,00158611
уг 0,00128586	ы 0,00320813	оц 8,975E-05
их 0,00174766	мн 0,00116021	це 0,00112268
х 0,00238733	аж 0,00124507	дс 0,00027741
ф 0,00078816	у 0,00386248	уш 0,00078816
фо 6,364E-05	ещ 0,00067393	ую 0,00104109
рм 0,00024477	ще 0,00123364	яю 8,975E-05
ма 0,00287524	оз 0,00090891	ющ 0,00045038
ах 0,00101335	зв 0,00078	щу 6,854E-05
пр 0,00619269	ит 0,00342842	жи 0,00114389
ри 0,00542738	еб 0,00109004	вь 0,00016971
ия 0,00164486	з 0,00416926	ье 0,00096113
ят 0,00144904	за 0,00566888	зм 0,00026762
ог 0,00481545	ащ 0,0002464	яя 0,00020887
го 0,00936002	щи 0,00067557	еш 0,00075063
ен 0,00699064	ищ 0,00010933	ож 0,00165791
во 0,00729742	ща 0,00042427	ыт 0,00055481
йн 0,0003541	ад 0,00220457	нр 4,08E-05
на 0,01068015	до 0,00365034	шь 0,00033615
и 0,01082049	ос 0,00674751	вр 0,00049117
м 0,00573742	ти 0,00349695	мя 0,00045038
ми 0,00359813	уж 0,00131034	ув 0,00087138
нр 0,00069841	жа 0,00139846	р 0,00072778
т 0,00684541	сы 0,00041448	ею 0,00035573
ом 0,00504064	ан 0,00532457	юс 0,00014849
м 0,0054437	нт 0,00061845	нг 0,00017297
ев 0,00138214	хр 7,669E-05	гл 0,00169218
мб 3,26E-06	ис 0,00290787	чь 0,00020234
цг 2,774E-05	рю 0,00021377	зе 0,00021703
фк 1,142E-05	о 0,01201823	шн 0,00034431

дф	1,305E-05	н	0,00343821	тм	4,895E-05
оа	1,63E-06	зн	0,00157959	йе	1,142E-05
хт	2,121E-05	ж	0,00014686	рк	0,00026435
мш	6,53E-06	г	0,00066741	вя	0,00013054
яш	3,26E-06	ов	0,00834178	тс	0,00084201
ао	4,9E-06	зд	0,00069188	х	0,00112594
кч	3,26E-06	тв	0,00219804	хо	0,00234653
лв	9,79E-06	ву	0,00082896	ыч	0,00013544
пг	1,469E-05	уй	9,464E-05	ря	0,0011031
кг	3,26E-06	йт	0,00028067	об	0,00316733
дг	9,79E-06	нж	0,00040306	ьт	0,00010607
нх	1,63E-06	жу	0,00019745	еп	0,00097908
уу	1,63E-06	пу	0,00071147	ьц	8,975E-05
бц	6,53E-06	са	0,00146536	цо	0,00042916
гш	6,53E-06	ди	0,0028426	сж	2,285E-05
мщ	1,63E-06	сь	0,00344147	ои	0,00080774
цл	4,9E-06	сс	0,00077837	мы	0,00064456
мю	6,364E-05	аз	0,00416436	еи	0,00017134
уэ	2,285E-05	зы	0,00055481	лн	0,00026109
лш	1,63E-06	ыв	0,00111615	эн	4,895E-05
шр	3,26E-06	ай	0,00079469	ту	0,00182762
эц	1,63E-06	ию	0,00030188	зи	0,00032636
цн	2,611E-05	юл	4,08E-05	на	7,343E-05
юк	3,26E-06	од	0,00463269	тк	0,000359
йр	5,874E-05	из	0,00172155	сд	0,0003182
нп	9,79E-06	вл	0,00097255	бщ	0,00016971
лм	4,9E-06	ло	0,00652232	бм	6,364E-05
эш	1,63E-06	вн	0,00140498	лю	0,00085833
эх	4,9E-06	ш	0,00101335	юд	0,00036063
юи	1,63E-06	фр	0,00049933	рж	0,00038184
ью	1,63E-06	ей	0,00239222	ьб	0,00051891
хп	1,63E-06	йл	0,00023498	бк	0,00025782
уц	1,63E-06	ин	0,00331909	ша	0,00142946
иэ	1,63E-06	им	0,00294214	оя	0,00064456
ёл	1,63E-06	мп	0,00023335	ян	0,00062824
нж	2,448E-05	иц	0,00148168	тж	1,63E-06
кх	3,26E-06	цы	0,00012239	зб	0,00015013
ьп	1,63E-06	фе	0,0001436	зг	0,00045038
пц	1,63E-06	ео	0,00025619	яч	0,00012402
фм	4,9E-06	еч	0,00093013	лж	0,00034594
бз	1,63E-06	жн	0,00106883	оп	0,00112105
йщ	1,63E-06	чи	0,00121733	пы	0,00020398
лз	6,53E-06	зя	0,00053197	ш	8,649E-05
юг	1,63E-06	не	0,00243791	аг	0,0009905
иу	1,63E-06	ех	0,00107862	бр	0,00122712

пф	1,63E-06	ха	0,00101172	су	0,00075063
кё	3,26E-06	вш	0,00129728	йш	8,485E-05
ё	1,63E-06	ег	0,00348716	юц	4,732E-05
пш	1,305E-05	ее	0,00150779	рь	0,00083548
ыц	1,63E-06	че	0,00307758	уб	0,00052707
шс	1,63E-06	аш	0,00101661	йц	3,753E-05
кж	1,142E-05	шл	0,00047485	ея	0,00038511
дп	9,954E-05	ля	0,00179009	мм	1,632E-05
щ	8,16E-06	ял	0,00091544	рч	8,322E-05
дд	2,121E-05	ьк	0,001139	ух	0,00037205
тё	1,63E-06	дн	0,00171013	йм	4,569E-05
ёр	4,9E-06	бы	0,00371562	сю	0,00013381
ьи	1,305E-05	ыл	0,00228126	кс	7,017E-05
п	9,79E-06	л	0,00687152	ню	0,00015665
йб	4,9E-06	гр	0,00140988	юю	3,753E-05
шп	5,711E-05	ип	0,00022519	оо	0,00016971
хм	5,711E-05	пп	8,322E-05	дл	0,00063804
вг	2,285E-05	пб	1,63E-06	сё	0,00069025
дз	4,9E-06	гд	0,00090239	ё	0,00057603
ьи	7,669E-05	ое	0,00219314	бъ	7,343E-05
жъ	3,59E-05	яв	0,00027577	ья	5,059E-05
рф	2,121E-05	ся	0,00316896	рд	0,00041611
ьв	4,9E-06	дк	0,0002154	нб	1,142E-05
юх	6,53E-06	ки	0,00261252	рг	0,0002872
вц	1,469E-05	пи	0,00078	ау	0,00014034
дю	1,469E-05	со	0,00255704	гв	7,833E-05
юр	6,854E-05	оч	0,00136745	ты	0,00160733
вч	3,1E-05	чк	0,00023335	дъ	0,00050586
шт	0,00016644	зо	0,00071799	ьб	0,00010933
вб	9,79E-06	ых	0,00089096	вд	0,00028557
тю	6,201E-05	ут	0,00145394	д	0,0008877
сф	1,795E-05	цв	1,958E-05	чн	0,00047812
тг	9,79E-06	мф	1,63E-06	вм	0,00014523
ьд	8,322E-05	фы	1,958E-05	вз	0,0005336
жю	2,121E-05	кд	1,958E-05	йч	5,711E-05
ьщ	8,16E-06	дм	9,138E-05	пя	0,00037042
ыж	2,937E-05	жо	3,59E-05	нм	3,26E-06
юш	7,506E-05	пт	5,059E-05	нс	0,00041611
бж	4,9E-06	чо	9,791E-05	чр	2,121E-05
бш	4,9E-06	аи	0,00012565	эм	1,142E-05
шв	1,632E-05	фи	0,00065762	ящ	0,00015176
хе	4,08E-05	жм	1,142E-05	бб	2,611E-05
нл	3,26E-06	вх	4,08E-05	лт	0,00030352
мт	1,795E-05	зт	1,142E-05	ц	0,00030515
са	9,79E-06	ьч	4,243E-05	фу	7,343E-05

сь	1,305E-05	цк	0,0001795	нк	0,00022845
вь	1,632E-05	хл	6,364E-05	яд	0,00074247
бч	2,611E-05	пч	8,16E-06	вк	0,00013054
нз	1,958E-05	жб	5,222E-05	щ	7,506E-05
нш	3,753E-05	сц	2,121E-05	рз	3,427E-05
гм	1,142E-05	сг	8,16E-06	зж	0,0001795
ьо	4,569E-05	уе	0,00015339	яе	9,954E-05
эк	5,548E-05	ун	1,795E-05	дя	0,00037695
оэ	2,121E-05	зъ	1,632E-05	кл	0,00037205
тэ	4,9E-06	лг	0,00015339	лч	0,00021866
э	1,63E-06	ьд	3,753E-05	зю	0,00011749
яб	3,427E-05	лп	0,0001077	сч	0,00033289
эй	1,142E-05	мл	7,833E-05	юч	6,038E-05
кц	1,142E-05	дч	2,285E-05	шо	0,00034268
ьф	6,53E-06	бх	4,895E-05	юб	0,00031657
сз	5,222E-05	пк	6,364E-05	лл	0,00011912
нв	9,79E-06	ыз	4,895E-05	яц	3,427E-05
кз	6,53E-06	вв	2,285E-05	бн	0,00026109
эв	1,63E-06	ю	6,038E-05	бв	3,916E-05
ьв	1,63E-06	юн	8,485E-05	шу	0,00029862
эр	5,222E-05	юм	2,611E-05	ют	0,00043243
эс	7,669E-05	мг	6,201E-05	рщ	6,69E-05
рз	1,63E-06	як	7,833E-05	б	0,00023008
эз	3,26E-06	дц	0,00016971	ец	0,0002823
чэ	1,63E-06	яф	1,63E-06	ыш	0,00054013
шр	1,63E-06	фф	4,9E-06	ур	0,00059398
эд	3,26E-06	сб	4,406E-05	зк	0,00016971
ыш	1,63E-06	яп	1,469E-05	хс	3,916E-05
вэ	1,63E-06	юж	6,53E-06	мч	2,774E-05
ыу	4,9E-06	лт	4,406E-05	ущ	0,00020071
жл	1,63E-06	дб	3,59E-05	тц	0,00012402
вш	6,53E-06	оф	0,0004569	яж	0,00053197
рр	4,406E-05	дж	1,305E-05	йд	0,00018766
ём	1,63E-06	тп	6,201E-05	ыг	0,00011586
уф	6,53E-06	ьг	4,243E-05	ц	0,00011912
тф	1,469E-05	тт	0,00010933	жк	8,812E-05
бг	1,795E-05	фл	9,628E-05	дъ	0,00029699
бю	6,53E-06	тъ	4,08E-05	ью	0,00017787
фт	6,53E-06	чл	3,753E-05	зу	0,00026925
лб	2,448E-05	эг	9,79E-06	яр	4,08E-05
жж	1,305E-05	хд	1,63E-06	сш	5,222E-05
лщ	1,63E-06	бь	4,895E-05	рб	0,00012728
еф	1,469E-05	лф	8,16E-06	зр	0,00015176
бт	8,16E-06	тщ	1,958E-05	нф	1,632E-05
гт	3,26E-06	лд	0,00034268	нщ	0,00015013



жг 2,448E-05	зц 3,26E-06	яг 0,00031004
нф 9,79E-06	фс 6,53E-06	ьм 0,0003541
аа 9,79E-06	зз 6,53E-06	ьп 0,00012891
мз 3,26E-06	бд 1,469E-05	тх 2,285E-05
фь 1,63E-06	гс 3,916E-05	рш 0,00024314
мц 3,59E-05	ьз 0,00013544	ях 0,00014523
пс 3,26E-06	хь 3,26E-06	дх 5,385E-05
йо 1,305E-05	ье 0,00023172	гч 2,285E-05
эп 9,79E-06	тч 0,00016808	рх 0,00013707
кш 6,53E-06	рц 9,954E-05	хн 0,00015502
лѐ 6,53E-06	уя 7,017E-05	вп 0,00031983
ѐз 1,63E-06	тз 6,53E-06	уо 1,305E-05
щъ 1,958E-05	пн 5,385E-05	бс 7,18E-05
ян 1,142E-05	тб 3,1E-05	гк 0,00010933
мк 7,017E-05	сх 0,00019092	мр 4,406E-05
ср 0,00026925	оу 5,059E-05	ох 0,00056134
ху 0,00012565	юз 2,774E-05	хв 0,00013381
жс 2,121E-05	зч 1,305E-05	кю 1,632E-05
зш 1,305E-05	рп 4,406E-05	яй 1,958E-05
пъ 0,00061193	еу 6,854E-05	йк 4,08E-05
жч 4,895E-05	гн 0,00029862	щн 2,937E-05
рл 6,364E-05	тд 0,00013544	кв 0,0001746
	ощ 0,00012891	мс 0,00013707

### Частоти біграмм без перетину

сп 0,00178498	па 0,00121496	зч 1,629E-05
ас 0,00406832	вл 0,00099347	ях 0,00015309
иб 0,00049185	ря 0,0010977	мф 3,26E-06
ч 0,00505853	лн 0,00028012	кд 1,303E-05
то 0,0134297	уз 0,00095438	жо 3,909E-05
с 0,01563813	иа 8,795E-05	рц 7,166E-05
ка 0,00797379	сд 0,00032898	бм 7,166E-05
ча 0,00222471	бщ 0,00016938	ьз 0,00014658
ли 0,00665785	об 0,00315955	вт 0,00021498
к 0,01084344	ну 0,00292177	ср 0,00029641
ни 0,00708781	жи 0,0011661	пы 0,00022475
гу 0,00083712	лю 0,00089575	йд 0,00017264
в 0,01512674	вш 0,0012573	рб 0,00013681
б 0,00756663	эн 4,886E-05	ью 0,00019544
ес 0,0034983	ту 0,00182733	рщ 5,863E-05
пл 0,00075894	зи 0,0002671	ш 8,143E-05

ат	0,00485333	тк	0,00042344	гк	0,00010098
но	0,00856335	лы	0,00124753	жу	0,00020521
й	0,00574582	бк	0,00028338	зт	9,77E-06
эл	7,492E-05	ша	0,00140714	хл	5,863E-05
ек	0,00101627	оя	0,00071334	гв	8,469E-05
тр	0,00285988	им	0,00292502	хв	0,00013681
он	0,00702918	зб	0,00014332	жб	4,886E-05
би	0,00084363	ян	0,00064168	цк	0,00015961
бл	0,00080129	ля	0,00173612	яр	3,909E-05
ио	0,00029641	мн	0,00129639	аи	0,00013355
те	0,00475235	лж	0,00036156	шу	7,166E-05
ке	0,00050488	ыт	0,00051139	сю	0,00012052
вс	0,00310092	ьн	0,00107164	уи	9,77E-06
е	0,01279128	оп	0,00108467	зъ	6,51E-06
кн	0,00254392	уд	0,00197065	рп	4,56E-05
иг	0,00044625	йш	7,166E-05	уя	6,84E-05
и	0,01507462	юц	4,56E-05	эм	1,629E-05
ав	0,00359276	рь	0,00086969	оу	5,212E-05
ра	0,00842329	ще	0,00128336	зр	0,00015309
эт	0,00235826	уб	0,00057328	юн	8,469E-05
а	0,01250138	вь	0,00015309	мг	6,189E-05
же	0,003811	ея	0,00037459	як	8,795E-05
га	0,00101301	гл	0,00180778	пк	7,492E-05
д	0,00691517	рч	8,143E-05	дц	0,00015635
ру	0,00326052	ят	0,00140062	рк	0,00023778
гн	0,0008762	кс	6,84E-05	фф	6,51E-06
х	0,00247878	ьт	0,00010749	фы	1,954E-05
фо	6,189E-05	юю	2,932E-05	яп	9,77E-06
рм	0,00023127	ьц	0,000114	юж	6,51E-06
ах	0,00101627	оо	0,00017264	юм	2,28E-05
пр	0,00628652	ё	0,00056351	хр	5,863E-05
ня	0,00163515	дл	0,00066774	ц	0,00012378
тн	0,00083386	аг	0,00099021	фи	0,00064168
ог	0,00474909	ща	0,00043322	тч	0,00015309
о	0,01788239	ья	6,515E-05	фл	0,00010749
чт	0,00367745	рт	0,00065145	пт	4,886E-05
ен	0,0069282	рд	0,00036481	эг	1,303E-05
во	0,00716273	рг	0,00034201	чл	1,629E-05
йн	0,00034527	ты	0,00167423	бъ	5,212E-05
ми	0,003697	ус	0,00146251	лф	3,26E-06
т	0,00681746	су	0,00077849	тп	6,189E-05
ом	0,00501293	дь	0,00052768	лд	0,00034853
ле	0,00378494	вд	0,00027687	тщ	1,629E-05
в	0,00557644	д	0,00093809	зц	3,26E-06
ко	0,00848518	чн	0,00044299	тт	0,00012378

ла	0,00657316	ул	0,00153743	фс	9,77E-06
ев	0,00132245	ыб	0,00049836	бх	3,583E-05
ич	0,00121822	нц	0,00063517	зз	9,77E-06
ол	0,00697055	ге	0,00056351	бд	1,303E-05
ст	0,01106167	вз	0,00058631	йк	4,234E-05
ой	0,00375237	йч	4,886E-05	кж	9,77E-06
	9,77E-06	пя	0,0003583	лп	0,00010098
лс	0,00146903	яс	0,00085992	щ	9,77E-06
м	0,00539077	нм	3,26E-06	дд	2,28E-05
ь	0,00769692	з	0,0010749	йм	4,234E-05
пе	0,00256021	фа	0,00016612	оф	0,00045927
рв	0,0002899	ци	0,00022149	ёр	3,26E-06
ая	0,00285988	бб	2,28E-05	лп	0,00011726
п	0,01546224	це	0,00107164	п	1,303E-05
ть	0,00514974	ца	0,00056025	ьд	3,257E-05
м	0,00588262	ач	0,00077197	ьч	3,583E-05
ер	0,00705524	фу	8,795E-05	йб	9,77E-06
ны	0,00335498	нк	0,00020521	шп	4,234E-05
яз	0,00175892	ию	0,00031921	дб	3,909E-05
г	0,00499665	нр	3,583E-05	сб	4,886E-05
уа	2,932E-05	ью	0,00038761	лг	0,00014658
и	0,01075223	ды	0,00070031	ьг	2,932E-05
л	0,0026905	зг	0,0004723	вг	2,932E-05
ук	0,00096415	щ	7,817E-05	ьн	6,84E-05
та	0,00554387	зя	0,00050162	ть	3,583E-05
н	0,01430917	нд	0,00104884	жъ	3,583E-05
бо	0,00234849	мя	0,00041042	цв	2,28E-05
ль	0,00346573	яе	0,00011726	рф	1,629E-05
ше	0,00218237	кл	0,00035504	вх	3,909E-05
к	0,00453737	лч	0,00025407	юх	6,51E-06
по	0,00894771	зю	0,00010423	чо	0,00010749
ме	0,00277845	ыв	0,0011205	ау	0,00017589
ья	0,0004267	сч	0,00036807	уе	0,00014658
ф	0,00083386	ух	0,00034853	дю	6,51E-06
ам	0,00263513	юч	5,863E-05	шт	0,00018892
ил	0,00514322	ьш	0,00041693	ю	5,537E-05
ии	0,00063191	лл	0,00013681	мл	7,492E-05
ап	0,0006938	яц	3,909E-05	тю	8,143E-05
ар	0,00289245	рж	0,00037459	сф	1,303E-05
не	0,00942653	аф	0,00064168	юр	6,515E-05
я	0,00141691	иш	0,00037784	тг	6,51E-06
ре	0,0061204	бв	3,909E-05	ьд	7,817E-05
ду	0,00162538	шу	0,00028664	жю	2,28E-05
еж	0,00095112	ют	0,00037133	ьж	2,28E-05
да	0,00427679	яд	0,00075894	ч	8,143E-05

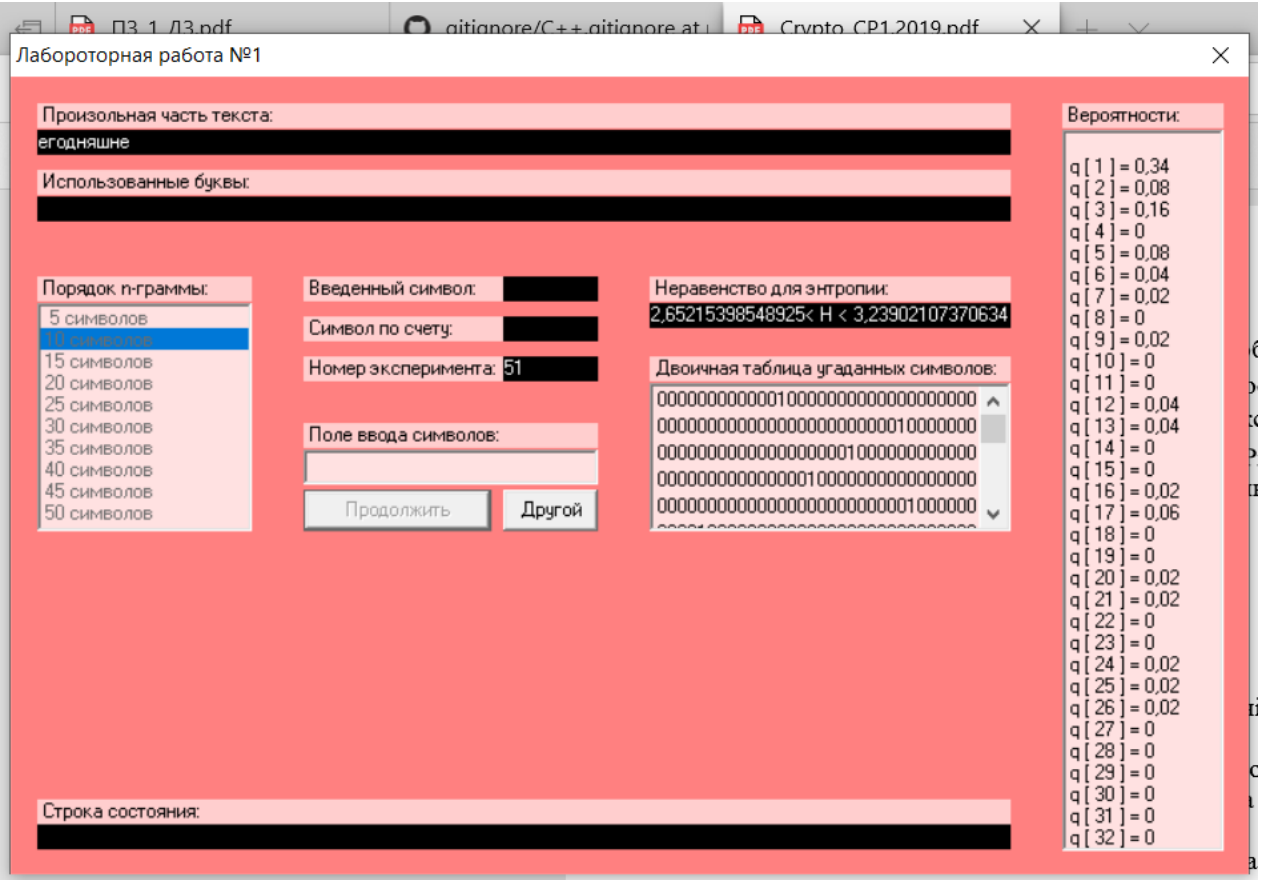
е	0,00390546	б	0,00020195	дм	6,84E-05
сл	0,00239409	сц	0,00028664	ыз	5,863E-05
вы	0,00267747	ыш	0,00063517	щъ	2,932E-05
ск	0,00489241	ур	0,00056676	шв	1,629E-05
аж	0,0012573	пп	3,583E-05	хе	4,56E-05
ет	0,00364814	дя	0,00040064	мт	1,954E-05
у	0,00388266	юд	0,0003583	въ	1,954E-05
на	0,01081086	ущ	0,00018241	ьб	8,143E-05
ы	0,00323772	тц	0,00013681	йц	2,932E-05
ещ	0,00061888	яж	0,00056351	бч	2,932E-05
зв	0,00074591	нс	0,0004267	нз	2,28E-05
ит	0,00337778	бр	0,00125405	гс	4,56E-05
се	0,00276216	ыг	0,00014332	юш	6,84E-05
бе	0,00203905	жк	0,00010423	вв	2,932E-05
з	0,00405204	дъ	0,00029967	сц	1,954E-05
аш	0,00025732	зо	0,00074266	гм	1,954E-05
иц	0,00010098	лу	0,00123776	хм	5,212E-05
до	0,00373934	зу	0,00028664	пч	3,26E-06
жа	0,00140062	сё	0,0007166	эк	7,166E-05
сы	0,00043973	оз	0,00094135	оэ	2,932E-05
э	0,00234523	ям	0,00030944	э	3,26E-06
го	0,00935487	вк	0,000114	эй	2,28E-05
а	0,00257975	сш	3,909E-05	хс	4,234E-05
нт	0,00068403	ха	0,0010293	ын	1,629E-05
их	0,00176544	иф	9,77E-06	нш	2,932E-05
ри	0,00528654	яг	0,00035504	сз	7,166E-05
я	0,00893468	шо	0,00036481	мч	1,629E-05
ве	0,00441685	эж	0,00015961	яб	3,583E-05
рю	0,00022475	ьм	0,0003583	эв	3,26E-06
ва	0,00602594	ып	0,00013029	ъв	3,26E-06
с	0,00279799	тх	1,954E-05	эр	7,817E-05
зн	0,00158629	рш	0,00023778	эс	9,12E-05
аю	0,0008534	еи	0,00017264	эз	6,51E-06
ж	0,00015961	дх	3,909E-05	чэ	3,26E-06
г	0,00062865	гч	1,954E-05	эд	3,26E-06
мо	0,00354065	бя	0,00046579	ыщ	3,26E-06
уж	0,00129639	рх	0,00013355	вэ	3,26E-06
рн	0,00065471	ут	0,00136479	вб	6,51E-06
ый	0,00158303	хн	0,00016938	жл	3,26E-06
р	0,0042833	яя	0,00020195	пн	4,886E-05
аб	0,00070031	пу	0,00075569	чр	2,932E-05
ов	0,00839723	ню	0,00017264	пр	4,56E-05
ор	0,00654059	нщ	0,00011726	дж	1,954E-05
др	0,00220191	бс	7,817E-05	бъ	5,863E-05
ву	0,00082735	мр	3,583E-05	ём	3,26E-06

йт	0,00032573	ящ	0,00017589	рз	3,909E-05
иж	0,0003811	нч	0,00019869	тф	1,629E-05
уг	0,00120193	чк	0,00023127	тб	2,28E-05
ад	0,00210745	вм	0,00015635	мк	5,863E-05
сс	0,00075569	яй	2,932E-05	бг	2,28E-05
зы	0,00061237	ай	0,00080454	уо	1,954E-05
юл	5,537E-05	шн	2,606E-05	бю	9,77E-06
од	0,00449828	чш	7,817E-05	ьв	6,51E-06
из	0,00170355	дк	0,00018892	лт	3,257E-05
ан	0,00549501	кв	0,00019544	лб	1,954E-05
ло	0,00644939	мс	0,00015635	еф	3,26E-06
вн	0,00139411	зш	9,77E-06	бг	6,51E-06
йл	0,00021172	дш	9,12E-05	гт	6,51E-06
ин	0,00318886	пъ	0,00068077	нв	3,26E-06
нн	0,00278822	бн	0,00024429	жг	2,606E-05
мп	0,00025407	жч	3,909E-05	мз	6,51E-06
цы	0,00012378	юб	0,00029967	фь	3,26E-06
ео	0,00027361	рл	4,234E-05	еа	9,77E-06
ро	0,00727999	зк	0,00017589	пс	6,51E-06
жн	0,00112376	зм	0,00025407	мц	5,212E-05
чи	0,00120845	еу	6,84E-05	дч	2,606E-05
си	0,00165143	уй	9,12E-05	лѐ	9,77E-06
ех	0,00098369	гн	0,0002671	мб	3,26E-06
ее	0,00148531	тд	0,00010098	цг	3,909E-05
еч	0,00090552	ощ	0,00011726	вч	2,932E-05
шл	0,00043647	ох	0,00060259	фк	1,629E-05
ял	0,00095764	жс	1,954E-05	ьо	4,56E-05
ьк	0,00115633	ц	0,00029641	дф	1,303E-05
дн	0,00171984	ье	0,00023127	аа	6,51E-06
ей	0,00250158	нр	0,00069705	оа	3,26E-06
у	0,00391523	зл	0,00019544	хт	1,303E-05
бы	0,00362208	тз	9,77E-06	эп	9,77E-06
л	0,00676208	яч	0,00013029	мш	3,26E-06
гр	0,00157	дт	0,00030618	жм	9,77E-06
ип	0,00029641	юс	0,00014332	ао	3,26E-06
ак	0,0044364	вп	0,00030944	вц	1,954E-05
о	0,01201279	яю	8,143E-05	яи	6,51E-06
пб	3,26E-06	сх	0,00018892	кг	6,51E-06
ыл	0,00234523	юз	3,583E-05	дг	1,303E-05
ое	0,00221494	ху	0,00010098	пг	6,51E-06
уп	0,00058631	сж	1,629E-05	кю	3,26E-06
от	0,00677185	бц	6,51E-06	сь	6,51E-06
яв	0,0002899	кц	3,26E-06	уу	3,26E-06
ед	0,00261233	иу	3,26E-06	кз	6,51E-06
ки	0,00267096	нп	9,77E-06	жж	9,77E-06

ис	0,00278496	пф	3,26E-06	гш	9,77E-06
оч	0,00138434	нл	3,26E-06	нф	6,51E-06
аз	0,00402598	кч	3,26E-06	уэ	9,77E-06
ос	0,00660573	сг	3,26E-06	нб	9,77E-06
ых	0,00079477	нж	1,954E-05	лш	3,26E-06
сн	0,00072637	кё	6,51E-06	мю	7,166E-05
ым	0,0015016	эх	6,51E-06	вщ	9,77E-06
пи	0,00076546	пш	9,77E-06	мм	3,26E-06
са	0,00158303	ыц	3,26E-06	цн	2,932E-05
ез	0,00112376	кш	3,26E-06	йр	7,166E-05
ф	6,51E-05	шъ	0,00033224	бш	6,51E-06
ня	0,00317258	вр	0,00053419	лм	9,77E-06
зь	0,00097718	ув	0,00086318	йо	1,303E-05
ид	0,00183384	ею	0,00036807	юн	3,26E-06
че	0,00310092	ик	0,00171984	лв	1,303E-05
уч	0,0007394	нг	0,00015309	ьо	3,26E-06
рс	0,00025732	йс	0,00060585	кх	6,51E-06
кт	0,00041367	за	0,00575885	бж	3,26E-06
ив	0,00275565	чь	0,00019544	ьп	3,26E-06
шк	0,00043973	т	0,00335173	фт	3,26E-06
ас	0,00089249	зе	0,00021824	уф	3,26E-06
бу	0,00118564	ма	0,00282731	фм	6,51E-06
нь	0,00111073	шн	0,00030944	ьу	3,26E-06
ви	0,00337778	ий	0,00147228	кк	9,77E-06
де	0,00450154	фе	0,00015635	ьф	3,26E-06
ын	0,0003583	йе	1,954E-05	ьщ	3,26E-06
еб	0,00109444	вя	0,00014006	лз	6,51E-06
жд	0,00065797	тс	0,00084363	юг	3,26E-06
мь	3,257E-05	х	0,00114656	хи	0,00018241
ю	0,00230289	тм	3,583E-05	кр	0,0017915
ся	0,00304554	ьч	0,00013681	ел	0,00625721
ш	0,00109118	ые	0,00082409	ьс	0,00095438
ок	0,00167098	щи	0,00071008	ац	9,446E-05
ие	0,00236477	хо	0,00241038	ры	0,001456
тв	0,00227357	ж	0,00169378	св	0,00179801
ал	0,0092376	еп	0,00091529	со	0,00263513
му	0,00237129	цо	0,00046579	оц	9,12E-05
тя	0,00051465	он	0,00080454	ку	0,00163515
сь	0,00341687	мы	0,00063517	дс	0,00025407
ою	0,00077197	ож	0,00165469	ую	0,00107815
ош	0,00129965	гд	0,00093809	ющ	0,00044299
ши	0,00202928	р	0,00065145	ьс	0,00077197
дв	0,00095112	тл	0,00030944	ег	0,00350807
ун	0,00032247	ыр	0,00036156	аш	0,00103907
ди	0,00297388	иц	0,00153743	зд	0,00069705

чу 0,00060585	н 0,0033615	ье 0,00091855
лк 0,00074591	фр 0,0004495	уш 0,00080454
ба 0,00117262	цу 0,00045927	см 0,00107164
шм 2,28E-05	зс 0,00012703	
ти 0,00342013	ык 0,00019869	ум 0,00107164
еш 0,00079477		ем 0,00382403

**Скріншоти результатів рожевої програми:**



H=2.956    R=1.71974





## Код програми

```
namespace test
{
    public partial class Form1 : Form
    {
        public string path    @"E:\Codes\test\text.txt";
        public char[] lang    new char[] { 'a', 'б', 'в', 'г', 'д', 'е', 'ё', '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', ' ' };
        public int[] count    new int[34] { 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, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0 };

        public string[] bigrams    new string[1024];
        public int[] bigramsCount  new int[1024];
        public int kolvo    0;
        public double summaBigram    0;
        public double spaceBigrams    0;

        public string[] nocrossbigrams    new string[1156];
        public int[] nocrossbigramsCount  new int[1156];
        public int nocrosskolvo    0;
        public double summaNocrossBigram    0;
        public double spaceNocrossBigrams    0;

        public double letters    0;

        public bool lettersCheck(char a)
        {
            for (int i    0; i < 34; i++)
            {
                if (lang[i]    a) return false;
            }
            return true;
        }

        public bool lettersCount(char a)
        {
            //Debug.WriteLine("AG");
            for (int i    0; i < 34; i++)
            {
                if (lang[i]    a)
                {
                    count[i]++;
                    letters++;
                    return false;
                }
            }
            return true;
        }

        public void check(char a, char b) ///с пересекающимися биграммami
        {
            if (lettersCount(b) || lettersCheck(a)) return;
            if (b    ' ' || a    ' ') spaceBigrams++;
            string z    a.ToString() + b.ToString();
            bool isadd    false;
            summaBigram++;
            for(int i    0; i < kolvo; i++)
            {
                //Debug.WriteLine(i);
                if (bigrams[i]    z)
            }
        }
    }
}
```

```

        {
            bigramsCount[i]++;
            isadd true;
            break;
        }

    }
    if (isadd false)
    {
        bigrams[kolvo] z;
        bigramsCount[kolvo] 1;
        kolvo++;
    }

}///
public void checkNocross(char a, char b)
{
    if (lettersCheck(b) || lettersCheck(a)) return;
    if (b ' ' || a ' ') spaceNocrossBigrams++;
    string z a.ToString() + b.ToString();
    bool isadd false;
    summaNocrossBigram++;
    for (int i 0; i < nocrosskolvo; i++)
    {
        if (nocrossbigrams[i] z)
        {
            nocrossbigramsCount[i]++;
            isadd true;
            break;
        }
    }
    if (isadd false)
    {
        nocrossbigrams[nocrosskolvo] z;
        nocrossbigramsCount[nocrosskolvo] 1;
        nocrosskolvo++;
    }
}
public double entropybigrams()
{
    double answ 0f;

    for (int i 0; i < kolvo; i++)
    {
        if (bigramsCount[i] != 0) answ answ + (bigramsCount[i] / summaBigram) *
Math.Log((bigramsCount[i] / summaBigram), 2);
    }

    answ -(answ / 2);
    return answ;
}
public double entropybigrams1()
{
    double answ 0f;

    for (int i 0; i < kolvo; i++)
    {
        //string bb bigrams[i];
        if (bigrams[i][0] != ' ' || bigrams[i][1] != ' ')
        {

```

```

        if (bigramsCount[i] != 0) answ = answ + (bigramsCount[i] /
(summaBigram-spaceBigrams)) * Math.Log((bigramsCount[i] / (summaBigram - spaceBigrams)),
2);
    }
}

    answ = -(answ/2);
    return answ;
}
public double entropy()
{
    double answ = 0f;

    for (int i = 0; i < 34; i++)
    {
        if(count[i]!= 0)answ = answ + (count[i] / letters) * Math.Log((count[i] /
letters), 2);
    }

    answ = -answ;
    return answ;
}

public double entropy1()
{
    double answ = 0f;

    for (int i = 0; i < 33; i++)
    {
        if (count[i] != 0) answ = -((count[i] / (letters-count[33])) *
Math.Log((count[i] / (letters-count[33])), 2));
    }

    return answ;
}

public double nocrossentropy()
{
    double answ = 0f;

    for (int i = 0; i < nocrosskolvo; i++)
    {
        if (nocrossbigramsCount[i] != 0) answ = answ +
(nocrossbigramsCount[i] / summaNocrossBigram) * Math.Log((nocrossbigramsCount[i] /
summaNocrossBigram), 2);
    }

    answ = -(answ / 2);
    return answ;
}
public double nocrossentropy1()
{
    double answ = 0f;

    for (int i = 0; i < nocrosskolvo; i++)
    {
        if (nocrossbigrams[i][0] != ' ' || nocrossbigrams[i][1] != ' ')
        {
            if (nocrossbigramsCount[i] != 0) answ = answ +
(nocrossbigramsCount[i] / (summaNocrossBigram-spaceNocrossBigrams)) *
Math.Log((nocrossbigramsCount[i] / (summaNocrossBigram - spaceNocrossBigrams)), 2);
        }
    }
}

```

```

        answ = -(answ / 2);
        return answ;
    }

```

```

public Form1()
{
    InitializeComponent();

    button1.Click += Button1_Click;
    button2.Click += Button2_Click;
    button3.Enabled = false;
    button4.Enabled = false;
    button5.Enabled = false;
    button6.Enabled = false;
    button7.Enabled = false;
    button8.Enabled = false;

    textBox1.Enabled = false;
    textBox2.Enabled = false;
    textBox3.Enabled = false;
    textBox4.Enabled = false;
    textBox5.Enabled = false;
    textBox6.Enabled = false;

}

```

```

private void Button1_Click(object sender, EventArgs e)
{

```

```

    using (StreamReader sr = new StreamReader(path, System.Text.Encoding.UTF8))
    {
        string line;
        while ((line = sr.ReadLine()) != null)
        {
            for (int i = 0; i < line.Length - 1; i++)
            {
                line = line.ToLower();
                check(line[i], line[i + 1]);
            }

            if(line.Length != 0) lettersCount(line[0]);
            //if(line.)

        }
        button3.Enabled = true;
        button4.Enabled = true;
        button1.Enabled = false;
        button5.Enabled = true;
        button6.Enabled = true;
    }
    textBox1.Enabled = true;
    textBox1.Text = Math.Round(entropy(), 3).ToString();
    textBox2.Enabled = true;
    textBox2.Text = Math.Round(entropy1(), 3).ToString();
    textBox3.Enabled = true;
    textBox3.Text = Math.Round(entropybigrams(), 3).ToString();
    textBox4.Enabled = true;

```

```

        textBox4.Text = Math.Round(entropybigrams1(), 3).ToString();
    }

    private void Button2_Click(object sender, EventArgs e)
    {
        using (StreamReader sr = new StreamReader(path, System.Text.Encoding.UTF8))
        {
            string line;
            while ((line = sr.ReadLine()) != null)
            {
                for (int i = 0; i < line.Length - 1; i++)
                {
                    line = line.ToLower();
                    checkNocross(line[i], line[i + 1]);
                    i++;
                }
            }
        }

        button2.Enabled = false;
        button7.Enabled = true;
        button8.Enabled = true;
        textBox5.Enabled = true;
        textBox5.Text = Math.Round(nocrossentropy(), 3).ToString();
        textBox6.Enabled = true;
        textBox6.Text = Math.Round(nocrossentropy1(), 3).ToString();
        Debug.WriteLine(nocrossentropy());
        Debug.WriteLine(nocrossentropy1());
    }
}
}

```

## Висновки:

Під час данного комп'ютерного практикуму, ми навчились визначати ентропію на символ джерела та його надлишковості. Порівняли різні моделі джерел відкритого тексту для наближеного визначення ентропії та набули практичних навичок оцінки ентропії на символ джерела.