



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»

Виконали:

студенти 3 курсу ФТІ

групи ФБ-73

Маковецький Андрій та Бадарак Оксана

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.

Ентропії:

Letters entropy with spaces: 4.363885193043505

Letters entropy without spaces: 4.45646766252117

Bigrams (step = 1) entropy with spaces: 3.9519384485559352

Bigrams (step = 1) entropy without spaces: 4.135418600084237

Bigrams (step = 2) entropy with spaces: 3.9506476208696535

Bigrams (step = 2) entropy without spaces: 4.13484307316223

Скріншоти результатів CoolPinkProgram:

10-грамми:

Лабораторная работа №1

Произвольная часть текста:
ли_вообще_ко_всем_людям_однако_они_всегда_были_согласны_в_том_что_не_следует

Использованные буквы:

Порядок n-граммы:
5 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

Неравенство для энтропии:
 $1.94568043870115 < H < 2.77173259480082$

Двоичная таблица угаданных символов:

01000000000000000000000000000000
10000000000000000000000000000000
10000000000000000000000000000000
10000000000000000000000000000000
00000100000000000000000000000000

Вероятности:

$q[1] = 0.5$
$q[2] = 0.12$
$q[3] = 0.04$
$q[4] = 0.04$
$q[5] = 0.04$
$q[6] = 0.04$
$q[7] = 0$
$q[8] = 0.04$
$q[9] = 0.02$
$q[10] = 0$
$q[11] = 0$
$q[12] = 0.04$
$q[13] = 0$
$q[14] = 0$
$q[15] = 0.02$
$q[16] = 0.02$
$q[17] = 0.02$
$q[18] = 0$
$q[19] = 0$
$q[20] = 0$
$q[21] = 0$
$q[22] = 0$
$q[23] = 0$
$q[24] = 0$
$q[25] = 0$
$q[26] = 0$
$q[27] = 0$
$q[28] = 0.02$
$q[29] = 0$
$q[30] = 0.02$
$q[31] = 0.02$
$q[32] = 0$

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

$$1.9457 < H < 2.7717$$

$$0.4505 < R < 0.6143$$

20-грамми:

Лабораторная работа №1

Произвольная часть текста:
кое_какие_важные_уроки_слушая_что_ссорающиеся_говорят_друг_другу_они_говорят

Использованные буквы:
а,

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

Введенный символ: о
Символ по счету: 2
Номер эксперимента: 50

Неравенство для энтропии:
 $1.31211312836284 < H < 2.04913533215$

Двоичная таблица угаданных символов:

10000000000000000000000000000000
01000000000000000000000000000000
00000000010000000000000000000000
01000000000000000000000000000000
10000000000000000000000000000000

Вероятности:

$q[1] = 0.6$
$q[2] = 0.18$
$q[3] = 0.04$
$q[4] = 0.04$
$q[5] = 0$
$q[6] = 0$
$q[7] = 0$
$q[8] = 0.02$
$q[9] = 0.02$
$q[10] = 0.02$
$q[11] = 0$
$q[12] = 0$
$q[13] = 0.02$
$q[14] = 0$
$q[15] = 0$
$q[16] = 0$
$q[17] = 0$
$q[18] = 0.02$
$q[19] = 0.02$
$q[20] = 0.02$
$q[21] = 0$
$q[22] = 0$
$q[23] = 0$
$q[24] = 0$
$q[25] = 0$
$q[26] = 0$
$q[27] = 0$
$q[28] = 0$
$q[29] = 0$
$q[30] = 0$
$q[31] = 0$
$q[32] = 0$

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

$$1.3121 < H < 2.0491$$

$$0.5938 < R < 0.7399$$

30-грамми:

Лабораторная работа №1

Произвольная часть текста:		
Использованные буквы:		
Порядок n-граммы:	Введенный символ:	Неравенство для энтропии:
<input type="radio"/> 5 символов	<input type="text" value="e"/>	<input type="text" value="1.12844488067691 < H < 1.80436057091161"/>
<input type="radio"/> 10 символов	Символ по счету:	
<input type="radio"/> 15 символов	<input type="text" value="1"/>	
<input type="radio"/> 20 символов	Номер эксперимента:	
<input checked="" type="radio"/> 25 символов	<input type="text" value="50"/>	
<input type="radio"/> 35 символов	Поле ввода символов:	Двоичная таблица угаданных символов:
<input type="radio"/> 40 символов	<input type="text" value="e"/>	<input type="button" value="↑"/> 10000000000000000000000000000000 10000000000000000000000000000000 01000000000000000000000000000000 10000000000000000000000000000000 10000000000000000000000000000000 <input type="button" value="↓"/>
<input type="radio"/> 45 символов	<input type="button" value="Продолжить"/>	<input type="button" value="Другой"/>
<input type="radio"/> 50 символов		

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

$$1.1284 < H < 1.8044$$

$$0.6423 < R < 0.7763$$

Частоти:

Монограми з пробілами:

а	0.166601	б	0.068713	в	0.01406	г	0.033279	д	0.014059	е	0.024983	ж	0.068283
з	0.009487	и	0.014029	й	0.055747	к	0.00885	л	0.028578	м	0.041089	н	0.026051
о	0.053724	п	0.096345	р	0.022408	с	0.034962	т	0.042934	у	0.053332	ф	0.024831
х	0.000351	ц	0.007338	ч	0.00258	ш	0.012529	щ	0.008351	ы	0.002955	ь	0.017367
э	0.018744	ю	0.002339	я	0.005367		0.019731						

Монограми без пробілів:

а	0.08245	б	0.016871	в	0.039932	г	0.016869	д	0.029977	е	0.081934	ж	0.011383
з	0.016834	и	0.066891	й	0.010619	к	0.034291	л	0.049303	м	0.031259	н	0.064464
о	0.115605	п	0.026887	р	0.041951	с	0.051517	т	0.063993	у	0.029795	ф	0.000421
х	0.008805	ц	0.003096	ч	0.015034	ш	0.010021	щ	0.003546	ы	0.020838	ь	0.022491
э	0.002807	ю	0.00644	я	0.023675								

Біграми з пробілами (крок 1)

_а	0.00304867	_б	0.00632648	_в	0.01306807	_г	0.00357985	_д	0.00720433	_е	0.00376286	_ж	0.00249072
_з	0.00516593	_и	0.01283894	_й	2.98E-06	_к	0.00905079	_л	0.00314538	_м	0.00756588	_н	0.01657501
_о	0.01045089	_п	0.01565401	_р	0.00428659	_с	0.01521211	_т	0.00783519	_у	0.00430593	_ф	0.00021723
_х	0.0017676	_ц	0.00042851	_ч	0.00522396	_ш	0.00129595	_щ	0.00012945	_э	0.0019774	_ю	0.00030055
я	0.00368995	а	0.02073067	аб	0.00083173	ав	0.0031558	аг	0.00079602	ад	0.00179141	ае	0.00172892
аж	0.00117543	аз	0.0028359	аи	0.00012498	ай	0.00059813	ак	0.00356497	ал	0.00725938	ам	0.00310372
ан	0.00235532	ао	6.70E-05	ап	0.00062193	ар	0.00213958	ас	0.00429998	ат	0.00472552	ау	9.22E-05
аф	8.93E-06	ах	0.00096712	ац	0.00021277	ач	0.00083024	аш	0.0008347	ащ	0.00024252	аю	0.00141795
ая	0.00220058	б_	0.00018152	ба	0.00069038	бв	6.55E-05	бг	2.98E-06	бд	7.44E-06	бе	0.00198484
бж	7.44E-06	бз	8.93E-06	би	0.00073055	бк	0.00013093	бл	0.00079899	бм	6.25E-05	бн	0.0004672
бо	0.00231663	бр	0.00106979	бс	0.00012796	бт	1.49E-06	бу	0.00116799	бх	3.27E-05	бш	8.93E-06
бщ	0.0001592	бы	0.00350843	бь	0.00015772	бэ	8.93E-06	бю	2.98E-06	бя	0.00035858	в_	0.00470022
ва	0.00517337	вб	1.49E-06	вв	1.49E-05	вг	2.98E-06	вд	0.00045083	ве	0.00437438	вз	0.00062789
ви	0.00254726	вк	0.00036304	вл	0.00067401	vm	0.00010862	вн	0.00108913	во	0.00600361	вп	0.00034816
вр	0.0005401	вс	0.00183902	вт	0.00012201	ву	0.00051927	vx	1.79E-05	вц	5.80E-05	вч	2.08E-05
вш	0.00052225	вщ	2.98E-06	вы	0.00255767	вь	0.00031692	вэ	1.49E-06	вя	0.00028121	г_	0.00072014
га	0.0009329	гв	7.44E-06	гг	1.49E-06	гд	0.000982	ге	8.03E-05	ги	0.00086892	гк	0.00019194
гл	0.0017676	гм	2.98E-06	гн	0.00036007	го	0.00679516	гр	0.00061152	гс	1.49E-06	гт	4.46E-05
гу	0.00065169	гч	3.57E-05	гш	1.49E-06	гэ	1.49E-06	д_	0.00106384	да	0.00414227	дб	3.42E-05
дв	0.00052671	дг	8.93E-06	дд	3.57E-05	де	0.00479991	дж	1.34E-05	дз	0.00010713	ди	0.0020384
дж	0.00031841	дл	0.00056986	дм	1.64E-05	дн	0.00152954	до	0.00418988	дп	4.61E-05	др	0.00101622
дс	0.00022765	дт	0.0002827	ду	0.00188515	дх	4.76E-05	дц	0.00013986	дч	6.70E-05	дш	3.87E-05
ды	0.00082131	дь	0.00065467	дэ	2.08E-05	дю	4.17E-05	дя	0.00029906	е_	0.0189958	еа	2.98E-05
еб	0.00116204	ев	0.00143283	ег	0.00226605	ед	0.00261421	ее	0.0016337	еж	0.00066806	ез	0.00112186
еи	7.89E-05	ей	0.00175421	ек	0.00095373	ел	0.0053043	ем	0.00323317	ен	0.00642468	ео	0.00013093
еп	0.00095224	ер	0.00541589	ес	0.00420922	ет	0.006078	еу	0.00012498	еф	2.08E-05	ех	0.00037346
ец	0.00030055	еч	0.00092844	еш	0.00090166	ещ	0.0007856	ею	0.00019938	ея	0.00018747	ж_	0.00043744
жа	0.00145069	жб	1.34E-05	жв	1.49E-06	жг	1.79E-05	жд	0.00077221	же	0.0036349	жж	1.64E-05
жи	0.0015102	жк	8.63E-05	жл	1.79E-05	жм	1.04E-05	жн	0.00094629	жо	9.08E-05	жр	3.42E-05
жс	2.23E-05	жт	1.49E-06	жу	0.00025145	жч	0.00013837	жь	3.27E-05	з_	0.0010564	за	0.00451721
зб	0.0001473	зв	0.00090463	зг	0.00030799	зд	0.0009091	зе	0.00039578	зж	6.40E-05	зз	4.46E-05
зи	0.00033031	зк	0.00023509	зл	0.00034816	зм	0.00031543	зн	0.00185688	зо	0.00074097	зп	1.79E-05
зр	0.00048803	зс	1.49E-06	зт	2.53E-05	зу	0.00042107	зч	1.04E-05	зш	4.46E-06	зы	0.00039875
зь	0.00015772	зю	2.98E-06	зя	0.00032733	и_	0.02010873	иа	5.65E-05	иб	0.00039578	ив	0.00255469
иг	0.0005282	ид	0.0010564	ие	0.00172446	иж	0.00027377	из	0.00180034	ии	0.00024401	ий	0.00077072
ик	0.0016203	ил	0.00469427	им	0.00212618	ин	0.0037941	ио	0.00054457	ип	0.00021277	ир	0.00091802
ис	0.00274961	ит	0.00473891	иу	1.49E-05	иф	2.08E-05	их	0.00132868	иц	0.00066657	ич	0.0007365
иш	0.00104896	ищ	0.0001592	ию	0.00012647	ия	0.00073204	й_	0.00760903	йг	1.49E-06	йд	0.00015474
йж	1.49E-06	йз	1.49E-06	йк	0.00010266	йл	2.98E-06	йм	4.76E-05	йн	0.00028419	йо	1.04E-05
йс	0.0002336	йт	0.00021128	йц	1.34E-05	йч	6.40E-05	йш	0.00010713	йю	2.98E-06	йя	1.49E-06
к_	0.00332691	ка	0.00649907	кв	0.00012349	ке	0.00047315	кж	4.46E-06	кз	1.49E-06	ки	0.00298469
кк	2.83E-05	кл	0.0007246	кн	0.00028419	ко	0.00908501	кп	2.98E-06	кр	0.00221248	кс	2.38E-05
кт	0.00056837	ку	0.0019402	кх	1.34E-05	кц	8.93E-06	кч	2.08E-05	кш	1.49E-05	кы	1.49E-06
кь	5.95E-06	кэ	0.00022616	кю	2.98E-06	л_	0.00335071	ла	0.00962809	лб	3.72E-05	лв	1.04E-05
лг	0.00021872	лд	0.00023211	ле	0.00338196	лж	0.00018003	лз	3.87E-05	ли	0.00732038	лк	0.000552

лл	3.87E-05	лм	4.46E-06	лн	0.00035412	ло	0.00615686	лп	2.23E-05	лс	0.00094629	лт	0.00010564
лу	0.00160394	лф	1.49E-06	лч	0.0001473	лш	7.44E-06	лщ	2.98E-06	лы	0.00083619	ль	0.00327185
лю	0.00098051	ля	0.00165899	м_	0.00676243	ма	0.00312753	мб	2.98E-05	мв	1.04E-05	мг	4.61E-05
мд	7.44E-06	ме	0.00387593	мз	2.98E-06	ми	0.00237764	мк	8.48E-05	мл	0.00021426	мм	1.93E-05
мн	0.00223034	мо	0.00359919	мп	0.00012498	мр	6.70E-05	мс	6.55E-05	мт	6.99E-05	му	0.00197144
мф	1.49E-06	мх	7.44E-06	мц	1.34E-05	мч	3.87E-05	мы	0.00075287	мь	0.00010415	мэ	2.53E-05
мя	0.00042107	н_	0.00267521	на	0.00988698	нб	1.49E-06	нг	0.00033924	нд	0.00010862	не	0.01099843
нж	5.95E-06	нз	1.79E-05	ни	0.00572984	нк	0.00040619	нм	1.64E-05	нн	0.00226753	но	0.01010421
нр	7.29E-05	нс	0.00018599	нт	0.00033031	ну	0.00300106	нф	1.64E-05	нх	1.49E-06	нц	0.00024401
нч	0.00012052	нщ	0.00046571	ны	0.00371227	нь	0.00102069	нэ	4.46E-06	ню	0.00019938	ня	0.00179141
о_	0.02247298	оа	1.49E-06	об	0.00352926	ов	0.00573132	ог	0.00418095	од	0.00454102	ое	0.00205625
ож	0.00236425	оз	0.00139712	ои	0.00106086	ой	0.003684	ок	0.00264397	ол	0.00577745	ом	0.00502458
он	0.00577745	оо	0.00017408	оп	0.00105342	ор	0.00585184	ос	0.00628482	от	0.00759118	оу	2.98E-05
оф	1.79E-05	ох	0.00065318	оц	9.37E-05	оч	0.00201459	ош	0.00102366	ощ	0.00028121	оэ	3.87E-05
ою	0.00023955	оя	0.00075436	п_	4.61E-05	па	0.00146408	пе	0.00178546	пи	0.00073799	пк	0.0001473
пл	0.00107425	пн	0.00011457	по	0.00921743	пп	1.19E-05	пр	0.00616727	пс	2.83E-05	пт	0.00011605
пу	0.00075733	пц	1.04E-05	пч	3.57E-05	пш	5.95E-06	пщ	1.49E-06	пы	0.00043	пь	6.10E-05
пя	0.00019491	р_	0.00059813	ра	0.0077117	рб	7.74E-05	рв	0.00030204	рг	0.00023211	рд	0.00014581
ре	0.00480735	рж	0.00036304	рз	8.18E-05	ри	0.00435354	рк	0.00029758	рл	0.00021426	рм	0.00022913
рн	0.00093885	ро	0.00705405	рп	7.74E-05	рр	4.91E-05	рс	0.00017111	рт	0.00045529	ру	0.00237912
рф	2.68E-05	рх	5.65E-05	рц	2.53E-05	рч	0.00012945	рш	0.00017706	рщ	5.80E-05	ры	0.00209494
рь	0.00047463	рэ	1.49E-06	рю	0.00012349	ря	0.00125577	с_	0.00226902	са	0.00125131	сб	6.10E-05
св	0.00113228	сг	5.80E-05	сд	0.00026633	се	0.00278532	сж	4.17E-05	сз	1.49E-06	си	0.00137183
ск	0.00273622	сл	0.00341915	см	0.00089422	сн	0.0010073	со	0.00263355	сп	0.00206072	ср	0.00018747
сс	0.00044636	ст	0.00985871	су	0.00083173	сф	2.98E-06	сх	0.00012796	сц	2.83E-05	сч	0.00040768
сш	7.29E-05	сщ	1.49E-06	сы	0.00044934	сь	0.00414375	сэ	2.98E-06	сю	0.00010564	ся	0.00427766
т_	0.00758969	та	0.00511236	тб	1.19E-05	тв	0.00177207	тг	4.46E-06	тд	0.00015474	те	0.00400092
тз	4.46E-06	ти	0.00297577	тк	0.0007856	тл	0.0002455	тм	6.25E-05	тн	0.00110252	то	0.01402478
тп	0.00012498	тр	0.00228241	тс	0.00202352	тт	5.65E-05	тү	0.00128553	тф	1.49E-06	тх	5.95E-06
тц	9.08E-05	тч	5.65E-05	тш	1.34E-05	тщ	7.44E-06	ты	0.00158311	ть	0.00722813	тэ	1.34E-05
тю	1.79E-05	тя	0.00069335	у_	0.0058697	уа	1.64E-05	уб	0.00059962	ув	0.00051332	уг	0.0009582
уд	0.00219314	уе	0.00017855	уж	0.00156823	уз	0.00022467	уй	1.64E-05	үй	0.00023062	үк	0.00094629
ул	0.00148937	ум	0.00121709	ун	0.00040322	уо	2.98E-06	уп	0.00082429	ур	0.00074692	үс	0.00122006
ут	0.00151913	уф	4.46E-06	ух	0.00063979	уц	1.04E-05	уч	0.00094481	уш	0.00113674	ущ	0.00027377
үз	1.34E-05	ую	0.0010073	үя	6.25E-05	ф_	2.98E-06	фа	4.46E-05	фе	1.79E-05	фи	7.74E-05
фл	1.93E-05	фо	0.00010415	фр	2.68E-05	фс	1.49E-06	фу	2.53E-05	фф	7.44E-06	фы	2.23E-05
фь	1.49E-06	х_	0.00293708	ха	0.00046422	хв	0.00027079	хг	1.49E-06	хе	0.00012052	хи	0.00044934
хк	1.49E-06	хл	0.00014581	хм	8.78E-05	хн	0.00029609	хо	0.0021738	хр	0.00025592	хс	4.46E-06
хт	1.19E-05	ху	0.00010118	хх	2.98E-06	хц	2.98E-06	хч	1.49E-06	хш	8.93E-06	ц_	0.00029014
ца	0.00050142	цв	0.00019342	це	0.00064128	цз	1.49E-06	ци	0.00020682	цк	7.44E-06	цо	0.00024401
цт	2.98E-06	цу	0.00028865	цц	1.49E-06	цы	0.00019789	цэ	1.49E-06	цю	1.49E-06	ч_	4.76E-05
ча	0.00159501	чв	1.49E-06	че	0.00284781	чи	0.00183902	чк	0.0004672	чл	2.23E-05	чм	1.49E-06
чн	0.00084958	чо	5.36E-05	чр	4.46E-06	чт	0.00353075	чу	0.00071716	чц	1.49E-06	чш	0.00013391
чь	0.00041661	ш_	7.29E-05	ша	0.00099539	шб	4.46E-06	шв	1.93E-05	ше	0.00198632	ши	0.00186283
шк	0.00124833	шл	0.00051332	шм	1.64E-05	шн	0.00031097	шо	0.00031989	шп	1.34E-05	шр	2.23E-05
шт	1.93E-05	шу	0.00026782	шц	4.46E-06	шь	0.00067252	шэ	1.49E-06	щ_	1.19E-05	ща	0.00037048
ще	0.00137927	щи	0.00092695	щн	4.31E-05	що	7.44E-06	щу	0.00019194	щь	2.38E-05	ы_	0.00483116
ыб	0.00036602	ыв	0.00104747	ыг	0.0001845	ыд	0.00012796	ые	0.00167982	ыж	5.06E-05	ыз	6.55E-05

ый	8.93E-06	ый	0.00171255	ык	0.00029906	ыл	0.00178248	ым	0.00126619	ын	0.00031989	ыо	1.49E-06
ып	0.00019789	ыр	0.00026038	ыс	0.00052373	ыт	0.00080048	ыу	1.49E-06	ых	0.00111889	ыц	2.98E-06
ыч	0.00017408	ыш	0.00050439	ыщ	1.93E-05	ыя	1.93E-05	ь_	0.01244763	ьб	7.59E-05	ьв	2.98E-06
ьг	0.00020533	ьд	2.08E-05	ье	0.00070228	ьз	0.00017706	ьи	0.00012052	ьк	0.00095522	ьм	0.00025443
ьн	0.00090314	ьо	4.61E-05	ьр	5.95E-06	ьс	0.00122155	ьт	2.38E-05	ьф	1.49E-06	ьц	0.00018003
ьч	0.00015474	ьш	0.00043446	ьщ	1.64E-05	ью	0.00046273	ья	0.0003318	э_	2.83E-05	эд	4.46E-06
эе	0.00022169	эй	3.27E-05	эк	1.49E-05	эл	1.49E-06	эм	1.19E-05	эн	1.49E-05	ео	1.49E-06
эп	2.98E-06	эр	2.98E-06	эс	7.44E-06	эт	0.00198335	эф	1.49E-06	эх	7.44E-06	эц	1.49E-06
ю_	0.00282549	юа	1.49E-06	юб	0.00033477	юв	1.34E-05	юг	5.95E-06	юд	0.00045827	юж	4.91E-05
юз	1.49E-05	юк	0.00013986	юм	7.44E-06	юн	0.00011308	юр	7.59E-05	юс	0.00032287	ют	0.0006874
юх	2.38E-05	юц	4.46E-06	юч	6.10E-05	юш	8.93E-06	ющ	0.00020235	юю	1.64E-05	я_	0.01254136
яб	1.93E-05	яв	0.00040322	яг	0.0001101	яд	0.00074394	яе	0.00022318	яж	0.00014879	яз	0.00028567
яи	1.34E-05	яй	6.40E-05	як	0.00019789	ял	0.00082429	ям	0.00034519	ян	0.00071121	яп	3.72E-05
яр	0.00015028	яс	0.00053266	ят	0.00153401	яу	2.98E-06	ях	0.00014432	яц	3.57E-05	яч	0.00016515
яш	4.46E-05	ящ	0.00014581	яю	0.00011457	яя	0.00019194						

Біграми без пробілів (крок 1)

аа	0.00047489	аб	0.00199242	ав	0.00568088	аг	0.00153894	ад	0.00307253	ае	0.00282259	аж	0.00164428
аз	0.00412052	аи	0.00222094	ай	0.0007177	ак	0.00561304	ал	0.00907121	ам	0.00491677	ан	0.00505424
ао	0.00161571	ап	0.00314216	ар	0.00320108	ас	0.00725375	ат	0.00694132	ау	0.0007177	аф	3.39E-05
ах	0.00136577	ац	0.00031779	ач	0.00182103	аш	0.00131221	ащ	0.00031065	аэ	0.00027851	аю	0.00175318
ая	0.00344388	ба	0.00083017	бб	8.93E-06	бв	8.57E-05	бг	7.14E-06	бд	1.07E-05	бе	0.00239233
бж	8.93E-06	бз	1.43E-05	би	0.00091408	бк	0.00016603	бл	0.0009605	бм	7.86E-05	бн	0.00057487
бо	0.00279759	бп	1.61E-05	бр	0.00128364	бс	0.00016961	бт	1.25E-05	бу	0.00140862	бх	3.93E-05
бч	8.93E-06	бш	1.25E-05	бщ	0.00019281	бы	0.00420978	бь	0.00018924	бэ	3.57E-05	бю	3.57E-06
бя	0.00043919	ва	0.00626647	вб	0.00014818	вв	0.00028744	вг	0.00024637	вд	0.00090159	ве	0.00534524
вж	0.00011069	вз	0.00088195	ви	0.00325821	vk	0.00092837	вл	0.00096586	vm	0.00045169	вн	0.00180674
во	0.00745192	вп	0.00092301	вр	0.00084981	вс	0.00274761	вт	0.00049275	ву	0.00071948	вф	3.57E-06
вх	0.00012497	вц	0.00010176	вч	0.00021245	вш	0.00068556	вщ	1.25E-05	вы	0.00306896	вь	0.00038027
вэ	0.00016603	вю	5.36E-06	вя	0.00038741	га	0.00113725	гб	2.50E-05	гв	6.96E-05	гг	1.25E-05
гд	0.00123723	ге	0.00011783	гж	1.61E-05	гз	3.04E-05	ги	0.00113725	гк	0.00026423	гл	0.00213881
гм	2.86E-05	гн	0.00053024	го	0.00821782	гп	8.39E-05	гр	0.00075162	гс	8.03E-05	гт	8.39E-05
гу	0.00081411	гф	1.79E-06	гх	1.07E-05	гц	1.79E-06	гч	5.18E-05	гш	7.14E-06	гэ	8.93E-06
гя	1.07E-05	да	0.0049864	дб	9.11E-05	дв	0.00069449	дг	3.04E-05	дд	6.43E-05	де	0.00580407
дж	3.04E-05	дз	0.00016603	ди	0.00255301	дк	0.00047489	дл	0.0007784	дм	7.14E-05	дн	0.00197099
до	0.00511137	дп	0.0001821	др	0.00124794	дс	0.00038741	дт	0.00039098	ду	0.00229056	дф	1.79E-06
дх	6.43E-05	дц	0.00017139	дч	0.00011426	дш	5.36E-05	дщ	1.79E-06	ды	0.0009855	дь	0.00078554
дэ	4.11E-05	дю	5.36E-05	дя	0.00037849	еа	0.00031779	еб	0.00242089	ев	0.00344209	ег	0.00317608
ед	0.00406517	ее	0.00231734	еж	0.00121045	ез	0.00226021	ei	0.00158715	ей	0.00210489	ек	0.00216738
ел	0.0068217	ем	0.00482215	ен	0.00976212	ео	0.00168534	еп	0.00363134	ер	0.00712342	ес	0.00767152
ет	0.00836422	еу	0.00084446	еф	6.43E-05	ех	0.00075698	ец	0.00040348	еч	0.00166927	еш	0.00121223
ещ	0.00095515	еэ	0.00019817	ею	0.00028922	ея	0.00058737	жа	0.00174247	жб	5.18E-05	жв	4.28E-05
жг	4.28E-05	жд	0.00093729	же	0.0043776	жж	2.32E-05	жз	1.79E-05	жи	0.00185138	жк	0.0001214
жл	2.50E-05	жм	3.75E-05	жн	0.00120152	жо	0.00014104	жп	4.82E-05	жр	5.36E-05	жс	6.07E-05
жт	3.21E-05	жу	0.00032493	жх	3.57E-06	жч	0.00017496	жш	1.79E-06	жь	3.93E-05	жэ	1.07E-05
жя	1.96E-05	за	0.00543094	зб	0.00021781	зв	0.00116938	зг	0.00040348	зд	0.00118367	зе	0.00048739
зж	9.11E-05	зз	8.93E-05	зи	0.00044633	зк	0.00039634	зл	0.00044454	зм	0.00045169	зн	0.00232984

зо	0.00093908	зп	0.0001214	зр	0.00063736	зс	0.00011248	зт	0.00016603	зу	0.00055345	зф	5.36E-06
зх	2.32E-05	зч	3.39E-05	зш	1.96E-05	зщ	1.79E-06	зы	0.00047847	зь	0.00018924	зэ	7.14E-06
зю	5.36E-06	зя	0.00039813	иа	0.0004499	иб	0.00137648	ив	0.00527205	иг	0.00120866	ид	0.00235841
ие	0.00268333	иж	0.00063914	из	0.00287079	ии	0.00186387	ий	0.0009248	ик	0.00340639	ил	0.00602188
им	0.00358849	ин	0.00694311	ио	0.00216381	ип	0.00267441	ир	0.0017514	ис	0.00548629	ит	0.00686991
иу	0.00055523	иф	4.82E-05	их	0.00183709	иц	0.00086052	ич	0.00147646	иш	0.00144075	ищ	0.00021067
из	0.00030172	ию	0.00018567	ия	0.00142111	йа	0.00021602	йб	0.0003035	йв	0.00064986	йг	0.00022852
йд	0.00063914	йе	0.00013747	йж	0.00021602	йз	0.0002553	йи	0.00079625	йй	1.79E-06	йк	0.00071234
йл	0.00019639	йм	0.00047489	йн	0.00113011	йо	0.00047311	йп	0.00079625	йр	0.00032671	йс	0.00123187
йт	0.00060165	йу	0.00023209	йф	3.04E-05	йх	9.11E-05	йц	5.53E-05	йч	0.00036063	йш	0.00022316
йщ	1.43E-05	йэ	5.89E-05	йю	2.50E-05	йя	0.00014104	ка	0.00786254	кб	0.00013211	кв	0.00042134
кг	0.00010533	кд	0.00016068	ке	0.00063557	кж	5.36E-05	кз	0.0001089	ки	0.00399197	кк	0.0002553
кл	0.000948	км	0.00015175	кн	0.00079447	ко	0.01114753	кп	0.00037849	кр	0.00275118	кс	0.00037849
кт	0.00088373	ку	0.00244589	кх	5.89E-05	кц	1.25E-05	кч	0.0001464	кш	3.93E-05	кы	1.79E-06
кь	7.14E-06	кэ	0.00033385	кю	7.14E-06	кя	7.68E-05	ла	0.0116617	лб	0.00022138	лв	0.00034278
лг	0.00035349	лд	0.00040705	ле	0.00419371	лж	0.0002678	лз	0.00014283	ли	0.00918547	лк	0.00085338
лл	1.00E-04	лм	0.00018032	лн	0.00078733	ло	0.00767866	лп	0.00033921	лр	0.00011248	лс	0.00150324
лт	0.00031422	лу	0.00202991	лф	5.36E-06	лх	3.21E-05	лц	8.93E-06	лч	0.00035171	лш	2.14E-05
лщ	3.57E-06	лы	0.00100335	ль	0.00392592	лэ	4.11E-05	лю	0.00118188	ля	0.00205312	ма	0.00395091
мб	0.00032493	мв	0.00059987	мг	0.00026066	мд	0.00045169	ме	0.00480965	мж	0.00011069	мз	0.00021959
ми	0.00359385	мк	0.00046061	мл	0.0003892	мм	0.00047311	мн	0.00344209	мо	0.00487571	мп	0.00081946
мр	0.0002678	мс	0.00081053	мт	0.00036242	му	0.00263513	мф	1.61E-05	мх	7.32E-05	мц	4.46E-05
мч	0.00032136	мш	7.14E-05	мщ	1.07E-05	мы	0.00090337	мь	0.00012497	мэ	0.00011248	мю	2.32E-05
мя	0.00069985	на	0.01194735	нб	0.00017853	нв	0.00026601	нг	0.00046954	нд	0.00024816	не	0.01324528
нж	4.28E-05	нз	0.00012854	ни	0.0071502	нк	0.00061236	нл	4.11E-05	нм	0.0001089	нн	0.00305647
но	0.0123187	нп	0.00039813	нр	0.00017853	нс	0.00053738	нт	0.00050346	ну	0.00374024	нф	2.50E-05
нх	4.64E-05	нц	0.00029279	нч	0.00020531	нш	1.79E-05	нщ	0.00056238	ны	0.00445437	нь	0.00122473
нэ	3.04E-05	ню	0.00023923	ня	0.00219416	оа	0.00036778	об	0.00555056	ов	0.00911227	ог	0.00541666
од	0.00676814	ое	0.00307789	ож	0.00330641	оз	0.00249231	ои	0.0030279	ой	0.00442045	ок	0.00439188
ол	0.00750727	ом	0.00726446	он	0.00966572	оо	0.00190136	оп	0.00364384	ор	0.0076608	ос	0.01004242
от	0.01044055	оу	0.00074983	оф	5.36E-05	ох	0.00108012	оц	0.00018924	оч	0.00331355	ош	0.00141219
ощ	0.00036242	оэ	0.00045169	ою	0.00031957	оя	0.00161393	па	0.00175854	пб	3.57E-06	пв	7.14E-06
пд	1.79E-06	пе	0.00214238	пз	3.57E-06	пи	0.00089266	пк	0.00017853	пл	0.001289	пн	0.00014104
по	0.01106362	пп	1.96E-05	пр	0.00740015	пс	3.93E-05	пт	0.00014104	пу	0.00090873	пх	3.57E-06
пц	1.25E-05	пч	4.64E-05	пш	7.14E-06	пщ	1.79E-06	пы	0.00051596	пь	7.32E-05	пя	0.00023566
ра	0.00926223	рб	9.82E-05	рв	0.00041955	рг	0.00029101	рд	0.00019103	ре	0.00577729	рж	0.0004374
рз	0.00011069	ри	0.00534524	рк	0.00041419	рл	0.00026958	рм	0.00029458	рн	0.0012033	ро	0.00851597
рп	0.00017675	рр	8.57E-05	рс	0.00026423	рт	0.00056059	ру	0.00287436	рф	3.21E-05	рх	7.50E-05
рц	3.04E-05	рч	0.00017675	рш	0.00021245	рщ	6.96E-05	ры	0.00251373	рь	0.00056952	рэ	5.36E-06
рю	0.00014818	ря	0.00152645	са	0.00152823	сб	0.00013747	св	0.00153359	сг	0.00013033	сд	0.00048918
се	0.00342603	сж	9.46E-05	сз	7.86E-05	си	0.00177282	ск	0.00351172	сл	0.00416336	см	0.00125686
сн	0.00153716	со	0.00332605	сп	0.00270297	ср	0.00029279	сс	0.00069985	ст	0.01204376	су	0.00106941
сф	1.25E-05	сх	0.00017675	сц	3.57E-05	сч	0.00053024	сш	0.00011426	сщ	1.79E-06	сы	0.00053917
сь	0.00497211	сэ	3.93E-05	сю	0.00013568	ся	0.00516493	та	0.00641287	тб	0.00024637	тв	0.00279224
тг	0.00024816	тд	0.00056952	те	0.00500961	тж	0.00011783	тз	0.00025709	ти	0.00431333	тк	0.00145682
тл	0.00051239	тм	0.00052488	тн	0.00218345	то	0.01746399	тп	0.00089266	тр	0.00296541	тс	0.00314573
тт	0.00055702	ту	0.00176747	тф	7.14E-06	тх	1.00E-04	тц	0.00012319	тч	0.00041598	тш	8.75E-05
тщ	1.43E-05	ты	0.00189958	ть	0.00867308	тэ	0.00016961	тю	4.28E-05	тя	0.00102299	ya	0.00020531

уб	0.00094622	ув	0.00113189	уг	0.00125865	уд	0.00290293	уе	0.00033028	уж	0.00196921	уз	0.00046061
уй	0.00070163	уй	0.00027672	ук	0.00159608	ул	0.00189244	ум	0.00185495	ун	0.00134613	уо	0.00041419
уп	0.00163178	ур	0.00104441	ус	0.0020067	ут	0.00209954	уу	0.00014461	уф	1.07E-05	ух	0.0008266
уц	2.32E-05	уч	0.00139076	уш	0.00140148	ущ	0.0003285	уэ	1.00E-04	ую	0.00123723	уя	0.00026244
фа	5.36E-05	фе	2.14E-05	фи	9.46E-05	фл	2.32E-05	фо	0.00012497	фп	1.79E-06	фр	3.21E-05
фс	1.79E-06	фу	3.04E-05	фф	8.93E-06	фы	2.68E-05	фь	1.79E-06	ха	0.00062665	хб	0.0001214
хв	0.00055702	хг	0.00011248	хд	0.00021424	хе	0.00021781	хж	6.78E-05	хз	9.82E-05	хи	0.0008516
хк	0.00026958	хл	0.00030172	хм	0.00027672	хн	0.00061594	хо	0.00280831	хп	0.00028387	хр	0.00041776
хс	0.00033028	хт	0.00014104	ху	0.00020353	хф	1.07E-05	хх	3.39E-05	хц	3.39E-05	хч	7.86E-05
хш	3.93E-05	хщ	5.36E-06	хэ	2.32E-05	хю	1.07E-05	хя	5.36E-05	ца	0.00061058	цб	5.36E-06
цв	0.0002428	цг	7.14E-06	цд	1.25E-05	це	0.00078376	цж	1.79E-06	цз	1.43E-05	ци	0.00028744
цк	4.46E-05	цл	3.57E-06	цм	1.79E-05	цн	3.21E-05	цо	0.00031957	цп	2.68E-05	цр	1.25E-05
цс	3.04E-05	цт	2.50E-05	цу	0.00036063	цх	3.57E-06	цц	1.79E-06	цч	7.14E-06	цы	0.00023745
цэ	3.57E-06	цю	1.79E-06	ця	1.79E-06	ча	0.00191743	чб	1.79E-06	чв	8.93E-06	че	0.0034171
чж	1.79E-06	чз	8.93E-06	чи	0.00220844	чк	0.00056238	чл	2.86E-05	чм	3.57E-06	чн	0.00102834
чо	6.61E-05	чп	3.57E-06	чр	5.36E-06	чс	7.14E-06	чт	0.00423835	чу	0.00086052	чц	1.79E-06
чч	1.79E-06	чш	0.00016068	чь	0.00049989	чя	1.79E-06	ша	0.00119973	шб	1.07E-05	шв	3.04E-05
шг	1.79E-06	шд	7.14E-06	ше	0.0023834	шз	5.36E-06	ши	0.00223879	шк	0.00150502	шл	0.00061594
шм	2.68E-05	шн	0.0003767	шо	0.0003892	шп	2.14E-05	шр	2.86E-05	шс	7.14E-06	шт	2.50E-05
шу	0.00032314	шц	5.36E-06	шч	1.79E-06	шш	1.79E-06	шь	0.00080696	шэ	5.36E-06	шы	3.57E-06
ща	0.00044633	щг	1.79E-06	ще	0.00165499	щи	0.00111404	щк	1.79E-06	щм	1.79E-06	щн	5.18E-05
що	8.93E-06	щс	3.57E-06	щт	1.79E-06	щу	0.00023031	щь	2.86E-05	ья	0.00013925	yb	0.0006445
ыв	0.00171926	ыг	0.000316	ыд	0.00036421	ые	0.00215667	ыж	0.00012854	ыз	0.00029815	ыи	0.00064271
ый	0.0020549	ык	0.00065878	ыл	0.00222986	ым	0.00171569	ын	0.00101763	ыо	0.00041598	ып	0.00078197
ыр	0.0004624	ыс	0.00111404	ыт	0.00119259	ыу	0.00016782	ыф	1.79E-06	ых	0.00144789	ыц	1.43E-05
ыч	0.00030529	ыш	0.00064093	ыщ	2.32E-05	ыэ	5.18E-05	ыю	3.57E-06	ыя	0.00012854	ба	0.00032671
ьб	0.00059987	ьв	0.0012408	ьг	0.00051239	ьд	0.00057309	ье	0.00129079	ьж	0.0002071	ьз	0.00060165
ьи	0.00142647	ьй	1.79E-06	ьк	0.00187994	ьл	0.00028208	ьм	0.00106226	ьн	0.00279224	ьо	0.00099799
ьп	0.00118545	ьр	0.00029458	ьс	0.00267976	ьт	0.00068556	ьу	0.00032671	ьф	1.79E-05	ьх	0.00015354
ьц	0.00025173	ьч	0.0009123	ьш	0.00061594	ьщ	3.21E-05	ьэ	0.00017496	ьо	0.00057844	ья	0.00078733
эд	8.93E-06	эе	0.0002678	эй	3.93E-05	эк	1.96E-05	эл	1.79E-06	эм	1.79E-05	эн	1.96E-05
эо	1.79E-06	эп	1.07E-05	эр	1.07E-05	эс	1.25E-05	эт	0.00237983	эф	1.79E-06	эх	8.93E-06
эц	3.57E-06	эш	1.79E-06	юа	6.61E-05	юб	0.00047847	юв	0.00022852	юг	9.46E-05	юд	0.00068378
юе	4.28E-05	юж	0.00015175	юз	9.82E-05	юи	0.00026601	юк	0.00038741	юл	5.89E-05	юм	0.00017853
юн	0.00043383	юо	0.00015889	юп	0.00031243	юр	0.00018389	юс	0.00074448	ют	0.00095515	юу	5.89E-05
юф	1.07E-05	юх	5.00E-05	юц	2.86E-05	юч	0.00030172	юш	5.36E-05	ющ	0.0002428	юэ	3.39E-05
юю	2.86E-05	юя	0.00010712	яа	0.00025352	яб	0.00053381	яв	0.00188351	яг	0.00039456	яд	0.00145325
яе	0.00056416	яж	0.00043383	яз	0.00099978	яи	0.00115689	яй	7.68E-05	як	0.00098014	ял	0.00122116
ям	0.00092658	ян	0.00263513	яо	0.00101049	яп	0.00163714	яр	0.00051774	яс	0.00195671	ят	0.00248695
яу	0.00040705	яф	1.79E-05	ях	0.00031779	яц	6.96E-05	яч	0.00060165	яш	0.00015354	ящ	0.00018746
яэ	0.00012319	яю	0.00016068	яя	0.00051417								

Біграми з пробілами (крок 2)

_а	0.00301445	_б	0.00635028	_в	0.01304873	_г	0.00361853	_д	0.00704959	_е	0.00379113	_ж	0.00251452
_з	0.00516295	_и	0.01307254	_к	0.00895705	_л	0.00324954	_м	0.00751083	_н	0.01681605	_о	0.0105967
_п	0.01548588	_р	0.00436545	_с	0.01522402	_т	0.00790958	_у	0.00428808	_ф	0.00019045	_х	0.0016575
_ц	0.00044041	_ч	0.00529389	_ш	0.00130934	_щ	0.0001101	_э	0.00190449	_ю	0.00027972	_я	0.00375244

а_	0.02049112	аб	0.00084214	ав	0.00316621	аг	0.00084809	ад	0.00177951	ае	0.0017557	аж	0.00128851
аз	0.00300255	аи	0.00012796	ай	0.00063681	ак	0.00357985	ал	0.00735312	ам	0.00305314	ан	0.00241335
ао	6.55E-05	ап	0.00062491	ар	0.00215445	ас	0.00443687	ат	0.00479991	ау	0.0001101	аф	5.95E-06
ах	0.0009582	ац	0.00019938	ач	0.00085702	аш	0.00082726	ащ	0.00025889	аю	0.00137778	ая	0.00219611
б_	0.00016367	ба	0.00063681	бв	5.36E-05	бд	1.19E-05	бе	0.00196698	бж	1.19E-05	бз	8.93E-06
би	0.00069335	бк	0.00011101	бл	0.00086297	бм	7.44E-05	бн	0.00048207	бо	0.00230027	бр	0.00109508
бс	0.00014284	бт	2.98E-06	бу	0.00109806	бх	4.17E-05	бш	1.19E-05	бщ	0.00016367	бы	0.00355902
бь	0.00016962	бэ	2.98E-06	бю	2.98E-06	бя	0.00038685	в_	0.00489811	ва	0.00508856	вв	8.93E-06
вг	2.98E-06	вд	0.0004791	ве	0.00432974	вз	0.00063979	ви	0.00257106	вк	0.00033924	вл	0.00067252
вм	7.74E-05	вн	0.00106235	во	0.00590392	вп	0.00037495	вр	0.00053861	вс	0.00172297	вт	0.0001101
ву	0.00053564	вх	2.08E-05	вц	6.55E-05	вч	1.49E-05	вш	0.00055052	вщ	5.95E-06	вы	0.00260082
вь	0.0002946	вя	0.0002827	г_	0.00069335	га	0.0009582	гв	8.93E-06	гд	0.00095224	ге	6.84E-05
ги	0.00085107	гк	0.00016069	гл	0.00175273	гм	5.95E-06	гн	0.00030055	го	0.00665679	гр	0.00060408
гт	5.36E-05	гу	0.00072311	гч	2.98E-05	гш	2.98E-06	д_	0.00101176	да	0.00421368	дб	4.46E-05
дв	0.00055052	дг	2.98E-06	дд	3.57E-05	де	0.00474635	дж	1.49E-05	дз	0.00010713	ди	0.00199376
дж	0.00035412	дл	0.00058325	дм	2.38E-05	дн	0.00149681	до	0.00418988	дп	5.06E-05	др	0.00104747
дс	0.00022616	дт	0.00024104	ду	0.00188961	дх	5.36E-05	дц	0.0001101	дч	8.33E-05	дш	3.87E-05
ды	0.00077965	дь	0.00056242	дэ	2.08E-05	дю	4.46E-05	дя	0.00024996	е_	0.01909846	еа	3.27E-05
еб	0.00116055	ев	0.00145515	ег	0.00233598	ед	0.00269902	ее	0.00164262	еж	0.00067252	ез	0.00112484
еи	8.93E-05	ей	0.00177653	ек	0.00094332	ел	0.00521354	ем	0.00318704	ен	0.00642468	ео	0.00013986
еп	0.00097308	ер	0.00543375	ес	0.00417202	ет	0.00610032	еу	0.00011605	еф	2.08E-05	ех	0.00036007
ец	0.00032733	еч	0.00096712	еш	0.00086892	ещ	0.00071716	ею	0.00017557	ея	0.00018747	ж_	0.00038387
жа	0.00143134	жб	1.19E-05	жг	2.68E-05	жд	0.00074394	же	0.00359472	жж	1.19E-05	жи	0.00153549
жк	8.33E-05	жл	1.79E-05	жм	1.79E-05	жн	0.00088678	жо	8.03E-05	жр	2.68E-05	жс	1.79E-05
жт	2.98E-06	жу	0.00024104	жч	0.00011903	жь	3.27E-05	з_	0.00104449	за	0.00445472	зб	0.00013391
зв	0.00091951	зг	0.00029758	зд	0.00085107	зе	0.00037792	зж	6.25E-05	зз	4.76E-05	зи	0.00031841
зк	0.00023211	зл	0.00033626	зм	0.00029163	зн	0.00184795	зо	0.00079453	зп	8.93E-06	зр	0.00049993
зс	2.98E-06	зт	1.49E-05	зу	0.00041066	зч	5.95E-06	зш	8.93E-06	зы	0.00036899	зь	0.00015176
зю	2.98E-06	зя	0.00034221	и_	0.0201251	иа	5.36E-05	иб	0.00044339	ив	0.00258594	иг	0.00055349
ид	0.00100283	ие	0.00181522	иж	0.00028567	из	0.00181522	ии	0.00022616	ий	0.00077667	ик	0.00156823
ил	0.00476122	им	0.00212767	ин	0.00380005	ио	0.00050588	ип	0.00020235	ир	0.00096117	ис	0.00265736
ит	0.00460946	иу	1.19E-05	иф	2.38E-05	их	0.001351	иц	0.00067847	ич	0.00072906	иш	0.00102961
ищ	0.00015772	ию	0.00011903	ия	0.00067847	й_	0.00756737	йг	2.98E-06	йд	0.00016069	йж	2.98E-06
йз	2.98E-06	йк	0.00010415	йл	5.95E-06	йм	3.87E-05	йн	0.00024401	йо	8.93E-06	йс	0.00022913
йт	0.00022021	йц	1.19E-05	йч	3.87E-05	йш	0.00013986	йю	2.98E-06	йя	2.98E-06	к_	0.00346082
ка	0.00643063	кв	0.00011903	ке	0.00051183	кж	5.95E-06	кз	2.98E-06	ки	0.00294898	кк	2.68E-05
кл	0.00072906	кн	0.00026484	ко	0.00933498	кп	5.95E-06	кр	0.00223778	кс	1.79E-05	кт	0.00052969
ку	0.00187771	кх	2.08E-05	кц	1.49E-05	кч	2.38E-05	кш	1.19E-05	кь	2.98E-06	кэ	0.00025889
кю	5.95E-06	л_	0.00334476	ла	0.00972778	лб	2.98E-05	лв	1.49E-05	лг	0.00022913	лд	0.00023211
ле	0.0031186	лж	0.00019045	лз	4.76E-05	ли	0.00727872	лк	0.00057135	лл	2.98E-05	лм	2.98E-06
лн	0.00035114	ло	0.00605866	лп	2.38E-05	лс	0.00091951	лт	0.00011605	лу	0.00164857	лф	2.98E-06
лч	0.00015474	лш	8.93E-06	лщ	2.98E-06	лы	0.00080048	ль	0.0034132	лю	0.00092844	ля	0.00169916
м_	0.00692758	ма	0.00315134	мб	2.68E-05	мв	1.19E-05	мг	4.76E-05	мд	5.95E-06	ме	0.00395479
мз	5.95E-06	ми	0.00230324	мк	7.74E-05	мл	0.00023509	мм	1.19E-05	мн	0.00217231	мо	0.00380005
мп	0.00013093	мр	6.84E-05	мс	7.74E-05	мт	8.93E-05	му	0.0020384	мф	2.98E-06	мх	2.98E-06
мц	1.49E-05	мч	4.76E-05	мы	0.00073501	мь	0.00010415	мэ	2.38E-05	мя	0.00038685	н_	0.00269902
на	0.00994203	нб	2.98E-06	нг	0.00035114	нд	0.00010415	не	0.01092403	нж	5.95E-06	нз	1.79E-05
ни	0.00563908	нк	0.00038387	нм	2.08E-05	нн	0.00233002	но	0.01026639	нр	7.44E-05	нс	0.0001845

нт	0.00033329	ну	0.00299064	нф	1.19E-05	нх	2.98E-06	нц	0.00025889	нч	0.00011605	нщ	0.000491
ны	0.00367209	нь	0.00103259	нэ	2.98E-06	ню	0.00017855	ня	0.00178843	о_	0.02233907	об	0.00344594
ов	0.00578191	ог	0.00422856	од	0.00471659	ое	0.00194317	ож	0.00238954	оз	0.0013629	ои	0.00111889
ой	0.00367805	ок	0.00271687	ол	0.00569264	ом	0.00482372	он	0.00575811	оо	0.00017855	оп	0.00111294
ор	0.00595748	ос	0.0062253	от	0.00738883	оу	3.27E-05	оф	2.08E-05	ох	0.00072014	оц	9.52E-05
оч	0.0020265	ош	0.00099093	ощ	0.0002827	оэ	5.36E-05	ою	0.00024996	оя	0.00080643	п_	5.06E-05
па	0.00143134	пе	0.00179439	пи	0.00072311	пк	0.00014284	пл	0.00109508	пн	0.00013986	по	0.00923678
пп	1.19E-05	пр	0.00625803	пс	3.87E-05	пт	0.00010415	пу	0.00073204	пц	1.79E-05	пч	2.08E-05
пш	5.95E-06	пщ	2.98E-06	пы	0.00042851	пь	5.06E-05	пя	0.0001964	р_	0.00056242	ра	0.00753166
рб	7.14E-05	рв	0.00028865	рг	0.00022318	рд	0.00015176	ре	0.00486835	рж	0.00033924	рз	7.14E-05
ри	0.00425832	рк	0.00033329	рл	0.00022616	рм	0.00022021	рн	0.00092546	ро	0.00697519	рп	6.25E-05
рр	5.36E-05	рс	0.00016367	рт	0.00044041	ру	0.00238656	рф	2.38E-05	рх	6.55E-05	рц	3.27E-05
рч	0.00012498	рш	0.00018747	рщ	7.74E-05	ры	0.00216933	рь	0.00041363	рю	0.00010713	ря	0.00119328
с_	0.00227051	са	0.00121114	сб	5.95E-05	св	0.00108318	сг	5.95E-05	сд	0.00024996	се	0.0028359
сж	2.98E-05	си	0.00139861	ск	0.00277341	сл	0.00338642	см	0.00088975	сн	0.00097903	со	0.00262165
сп	0.00208899	ср	0.00017855	сс	0.00041958	ст	0.01016819	су	0.00080346	сф	2.98E-06	сх	0.00012201
сц	2.98E-05	сч	0.00042851	сш	6.55E-05	сы	0.00045232	сь	0.00421368	сэ	2.98E-06	сю	9.52E-05
ся	0.00424344	т_	0.00754654	та	0.00496953	тб	5.95E-06	тв	0.00176463	тг	5.95E-06	тд	0.00016367
те	0.00395777	тз	5.95E-06	ти	0.00306801	тк	0.00078858	тл	0.00027675	тм	7.44E-05	тн	0.00103259
то	0.01395932	тп	0.00012796	тр	0.00229432	тс	0.00205625	тт	5.95E-05	ту	0.00129148	тф	2.98E-06
тх	1.19E-05	тц	9.22E-05	тч	6.25E-05	тш	1.19E-05	тщ	5.95E-06	ты	0.00161882	ть	0.00715969
тэ	1.19E-05	тю	1.79E-05	тя	0.00066062	у_	0.00583548	уа	1.49E-05	уб	0.00060408	ув	0.00052671
уг	0.0009701	уд	0.00236276	уе	0.00017855	уж	0.00162179	уз	0.00024699	уи	1.49E-05	уй	0.00021723
ук	0.00087785	ул	0.00149383	ум	0.00123792	ун	0.0003928	уо	5.95E-06	уп	0.00081238	ур	0.00071716
ус	0.00119328	ут	0.00150574	уф	5.95E-06	ух	0.0006755	уц	5.95E-06	уч	0.00087785	уш	0.00113972
ущ	0.00025294	уэ	2.08E-05	ую	0.00100878	уя	6.25E-05	ф_	2.98E-06	фа	5.06E-05	фе	1.49E-05
фи	9.22E-05	фл	1.79E-05	фо	0.0001101	фр	2.68E-05	фү	2.68E-05	фф	8.93E-06	фы	2.38E-05
х_	0.0028002	ха	0.00045827	хв	0.00029163	хе	0.00012796	хи	0.00047315	хл	0.00014581	хм	8.63E-05
хн	0.00031543	хо	0.00212172	хр	0.00024996	хс	2.98E-06	хт	1.79E-05	ху	8.93E-05	хх	5.95E-06
хц	2.98E-06	хч	2.98E-06	хш	5.95E-06	ц_	0.00027079	ца	0.00045529	цв	0.00017855	це	0.00068145
ци	0.00017259	цк	2.98E-06	цо	0.00023806	цт	2.98E-06	цу	0.00030948	цы	0.00021128	ч_	4.46E-05
ча	0.00160394	че	0.00285078	чи	0.00194317	чк	0.00042256	чл	2.08E-05	чм	2.98E-06	чн	0.00080346
чо	3.87E-05	чр	2.98E-06	чт	0.00355902	чу	0.00069038	чш	0.00012201	чь	0.00039875	ш_	6.55E-05
ша	0.00100878	шб	2.98E-06	шв	1.49E-05	ше	0.00196996	ши	0.00190151	шк	0.00124387	шл	0.00047612
шм	2.38E-05	шн	0.00033329	шо	0.00031841	шп	1.19E-05	шр	2.68E-05	шт	2.38E-05	шу	0.00025294
шц	2.98E-06	шь	0.00070228	шэ	2.98E-06	щ_	1.79E-05	ща	0.00034519	ще	0.00148491	щи	0.0008719
щн	4.17E-05	що	2.98E-06	щу	0.0001845	щъ	2.68E-05	ы_	0.00463327	ыб	0.00039578	ыв	0.00108615
ыг	0.00018747	ыд	0.00010118	ые	0.00170809	ыж	6.55E-05	ыз	7.44E-05	ыи	5.95E-06	ый	0.00173785
ык	0.00025592	ыл	0.00182117	ым	0.00123197	ын	0.00032138	ыо	2.98E-06	ып	0.0001845	ыр	0.00026484
ыс	0.00053266	ыт	0.00079155	ых	0.00119626	ыч	0.00016664	ыш	0.0005029	ыщ	2.68E-05	ыя	1.79E-05
ь_	0.01256963	ьб	8.93E-05	ьв	5.95E-06	ьг	0.00020235	ьд	2.08E-05	ье	0.00072609	ьз	0.00016367
ьи	0.00013093	ьк	0.00090463	ьм	0.00025294	ьн	0.00083024	ьо	3.87E-05	ьр	5.95E-06	ьс	0.00120221
ьт	1.79E-05	ьф	2.98E-06	ьц	0.00018152	ьч	0.00011308	ьш	0.00043149	ьщ	5.95E-06	ью	0.00053266
ья	0.00033031	э_	3.27E-05	эд	5.95E-06	эе	0.00019342	эй	3.87E-05	эк	2.08E-05	эм	2.08E-05
эн	1.49E-05	эп	2.98E-06	эр	2.98E-06	эс	5.95E-06	эт	0.00202352	эх	8.93E-06	ю_	0.002824
юб	0.00033329	юв	1.79E-05	юд	0.00052969	юж	4.76E-05	юз	2.08E-05	юк	0.00015474	юм	5.95E-06
юн	0.00012796	юр	6.25E-05	юс	0.00030948	ют	0.0006874	юх	2.98E-05	юц	5.95E-06	юч	6.55E-05
юш	8.93E-06	ющ	0.00021426	юю	1.49E-05	я_	0.01246251	яб	1.19E-05	яв	0.00038387	яг	0.00010118

яд	0.00078858	яе	0.00021723	яж	0.00014581	яз	0.00029758	яи	1.49E-05	яй	5.36E-05	як	0.00018152
ял	0.00077667	ям	0.00033329	ян	0.00071418	яп	3.57E-05	яр	0.00016664	яс	0.00052373	ят	0.0016337
яу	2.98E-06	ях	0.00017259	яц	4.76E-05	яч	0.0001845	яш	3.87E-05	ящ	0.00015772	яю	0.00012796
яя	0.00021128												

Біграми без пробілів (крок 2)

аа	0.00047132	аб	0.00197813	ав	0.00584156	аг	0.00157822	ад	0.00297434	ае	0.00287436	аж	0.00172105
аз	0.00409552	аи	0.00224236	ай	0.00076055	ак	0.00544522	ал	0.00904443	ам	0.00488106	ан	0.00502389
ао	0.00162821	ап	0.00312788	ар	0.00330998	ас	0.00735551	ат	0.00686277	ау	0.00066771	аф	3.57E-05
ах	0.00126758	ац	0.00027137	ач	0.00180674	аш	0.00129971	ащ	0.00029636	аэ	0.00024994	аю	0.00176747
ая	0.00348137	ба	0.00081053	бб	7.14E-06	бв	6.78E-05	бг	3.57E-06	бд	1.07E-05	бе	0.00249945
бж	1.79E-05	бз	2.14E-05	би	0.00091408	бк	0.0001464	бл	0.00097835	бм	7.14E-05	бн	0.00056416
бо	0.00277796	бп	1.43E-05	бр	0.00130685	бс	0.0001821	бт	1.07E-05	бу	0.0014104	бх	5.00E-05
бч	1.43E-05	бш	1.07E-05	бщ	0.00020353	бы	0.00424192	бь	0.00020353	бэ	2.50E-05	бю	3.57E-06
бя	0.0004499	ва	0.00643786	вб	0.00013568	вв	0.00029636	вг	0.0002428	вд	0.00088909	ве	0.00535953
вж	9.28E-05	вз	0.00085695	ви	0.00321	vk	0.00087124	вл	0.00096407	vm	0.00042491	vn	0.00189958
во	0.0076126	вп	0.00084981	вр	0.00079625	вс	0.00273511	вт	0.00047847	ву	0.00073198	vx	0.00011783
вц	0.00010355	вч	0.00022852	вш	0.00064986	вщ	1.43E-05	вы	0.00309217	вь	0.00038206	вэ	0.00015711
вю	3.57E-06	вя	0.00042491	га	0.0011069	гб	2.14E-05	гв	4.28E-05	гг	1.07E-05	гд	0.00126044
ге	0.00011426	гж	1.43E-05	гз	2.50E-05	ги	0.00106048	гк	0.00029279	гл	0.00217809	гм	2.50E-05
гн	0.0005963	го	0.0082089	гп	7.86E-05	гр	0.00077483	гс	7.50E-05	гт	8.57E-05	гу	0.00087124
гф	3.57E-06	гх	1.43E-05	гц	3.57E-06	гч	5.36E-05	гш	1.07E-05	гэ	1.43E-05	гя	1.07E-05
да	0.00510601	дб	7.50E-05	дв	0.00067485	дг	3.21E-05	дд	6.43E-05	де	0.00594511	дж	2.86E-05
дз	0.00016782	ди	0.00254944	дк	0.00043562	дл	0.00077126	дм	6.07E-05	дн	0.00206026	до	0.00504531
дп	0.00019281	др	0.00136398	дс	0.00040348	дт	0.00041776	ду	0.00227093	дф	3.57E-06	дх	7.86E-05
дц	0.0001464	дч	0.00011783	дш	5.71E-05	дщ	3.57E-06	ды	0.00100335	дь	0.00085695	дэ	3.21E-05
дю	5.00E-05	дя	0.00040348	еа	0.00031779	еб	0.00241732	ев	0.00331355	ег	0.00315644	ед	0.00398483
ее	0.00236019	еж	0.0011676	ез	0.00223522	ei	0.00166035	ей	0.00214595	ек	0.00214595	ел	0.00673779
ем	0.00476323	ен	0.00966572	ео	0.00166035	еп	0.00365633	ер	0.00720198	ес	0.00774471	ет	0.00838743
еу	0.00081768	еф	5.71E-05	ех	0.00068199	ец	0.00035349	еч	0.00171034	еш	0.00117474	ещ	0.00094979
ез	0.00017496	ею	0.00029636	ея	0.00064271	жа	0.00171391	жб	4.64E-05	жв	3.21E-05	жг	4.28E-05
жд	0.00096764	же	0.00442402	жж	2.14E-05	жз	1.07E-05	жи	0.00181745	жк	7.86E-05	жл	3.93E-05
жм	3.93E-05	жн	0.00117831	жо	0.0001464	жп	5.71E-05	жр	5.00E-05	жс	6.07E-05	жт	3.57E-05
жу	0.00033207	жх	7.14E-06	жч	0.00018567	жь	3.57E-05	жэ	7.14E-06	жя	1.79E-05	за	0.00535239
зб	0.00022138	зв	0.00121402	зг	0.00036778	зд	0.0011069	зе	0.00051417	эж	7.50E-05	зз	7.86E-05
зи	0.00041776	зк	0.00041776	зл	0.00047489	зм	0.00049275	зн	0.00243875	зо	0.00096407	зп	0.00011069
зр	0.00069985	зс	0.00011783	зт	0.00018567	зу	0.00047847	зф	3.57E-06	зх	1.07E-05	зч	3.93E-05
зш	1.43E-05	зщ	3.57E-06	зы	0.00051417	зь	0.00019281	зэ	1.07E-05	зю	7.14E-06	зя	0.00040348
иа	0.0004499	иб	0.00131399	ив	0.00538095	иг	0.00117831	ид	0.0023959	ие	0.00270297	иж	0.00061772
из	0.00274582	ии	0.00187816	ий	0.00094265	ик	0.00349565	ил	0.00590941	им	0.00354564	ин	0.00680564
ио	0.00214238	ип	0.002578	ир	0.00174604	ис	0.00560947	ит	0.00691276	иу	0.0005963	иф	4.28E-05
их	0.00181031	иц	0.00083553	ич	0.00146396	иш	0.00150681	ищ	0.00021781	иэ	0.0003035	ию	0.00020353
ия	0.0014711	йа	0.00020353	йб	0.0003285	йв	0.00062843	йг	0.0002428	йд	0.00059273	йе	0.00012497
йж	0.00019639	йз	0.00023566	йи	0.00080339	йй	3.57E-06	йк	0.00072841	йл	0.00019996	йм	0.00044276
йн	0.00112475	йо	0.00051774	йп	0.00081411	йр	0.00030707	йс	0.00126044	йт	0.00064271	йу	0.00026066
йф	2.86E-05	йх	8.57E-05	йц	6.78E-05	йч	0.00034992	йш	0.00025709	йщ	1.79E-05	йэ	5.00E-05
йю	1.79E-05	йя	0.00015711	ка	0.0079411	кб	0.00014997	кв	0.00041776	кг	0.00012854	кд	0.00017496

ке	0.00067485	кж	5.36E-05	кз	0.00011069	ки	0.00410266	кк	0.00027494	кл	0.00093194	км	0.00017139
кн	0.00076769	ко	0.01130821	кп	0.00037849	кр	0.00276367	кс	0.00039634	кт	0.00090337	ку	0.00239233
кх	7.86E-05	кц	7.14E-06	кч	0.00014283	кш	4.28E-05	кь	3.57E-06	кэ	0.00029993	кю	7.14E-06
кя	8.93E-05	ла	0.01181167	лб	0.00022495	лв	0.00036778	лг	0.00036421	лд	0.00036421	ле	0.00420621
лж	0.00028922	лз	0.00012854	ли	0.00919796	лк	0.00086767	лл	0.00011069	лм	0.0001464	лн	0.00079625
ло	0.00779113	лп	0.00035706	лр	0.00012497	лс	0.00158536	лт	0.00029993	лу	0.00203169	лф	3.57E-06
лх	2.50E-05	лц	7.14E-06	лч	0.00037135	лш	1.79E-05	лщ	3.57E-06	лы	0.00100692	ль	0.00383844
лэ	5.00E-05	лю	0.00124615	ля	0.00209954	ма	0.00382772	мб	0.00032136	мв	0.00058916	мг	0.0002428
мд	0.00044276	ме	0.00478465	мж	0.00012854	мз	0.00024637	ми	0.00368133	мк	0.00042848	мл	0.00041419
мм	0.00047847	мн	0.00335997	мо	0.00498104	мп	0.00086409	мр	0.00025352	мс	0.00077126	мт	0.00038563
му	0.0026637	мф	1.43E-05	мх	7.14E-05	мц	2.50E-05	мч	0.00031779	мш	6.43E-05	мщ	1.43E-05
мы	0.00093908	мь	0.00010355	мэ	1.00E-04	мю	2.50E-05	мя	0.00071056	на	0.01196878	нб	0.00017139
нв	0.00027851	нг	0.00046418	нд	0.0002678	не	0.01323635	нж	3.57E-05	нз	0.00012854	ни	0.00699488
нк	0.00060701	нл	5.00E-05	нм	0.00012497	нн	0.00303504	но	0.01227585	нп	0.00038563	нр	0.00018567
нс	0.00052845	нт	0.00048204	ну	0.0037206	нф	3.93E-05	нх	4.28E-05	нц	0.00029279	нч	0.00019281
нш	1.43E-05	нщ	0.00056059	ны	0.00442759	нь	0.00121759	нэ	3.57E-05	ню	0.00025352	ня	0.00226378
оа	0.00033921	об	0.00545951	ов	0.0092301	ог	0.00535953	од	0.00660925	ое	0.00292078	ож	0.00338854
оз	0.00259585	ои	0.00304218	ой	0.00436332	ок	0.00430619	ол	0.0071627	ом	0.00735194	он	0.01002635
оо	0.0018603	оп	0.00357778	ор	0.00761974	ос	0.0099228	от	0.01032628	оу	0.0007177	оф	5.00E-05
ох	0.00107833	оц	0.00017853	оч	0.00316716	ош	0.00134613	ощ	0.00035349	оэ	0.00040705	ою	0.00029993
оя	0.00157822	па	0.00175675	пв	3.57E-06	пд	3.57E-06	пе	0.00216024	пз	3.57E-06	пи	0.00084981
пк	0.0001821	пл	0.00125686	пн	0.00012854	по	0.01118681	пп	2.50E-05	пр	0.00747692	пс	3.93E-05
пт	0.0001464	пу	0.00094622	пц	1.43E-05	пч	5.36E-05	пш	3.57E-06	пщ	3.57E-06	пы	0.00054274
пъ	8.57E-05	пя	0.00025352	ра	0.00905157	рб	9.64E-05	рв	0.00042491	рг	0.00027137	рд	0.00021067
ре	0.00576301	рж	0.00042491	рз	0.00010712	ри	0.00531668	рк	0.0003892	рл	0.00023566	рм	0.0002678
рн	0.00119973	ро	0.00857667	рп	0.00016782	рр	6.43E-05	рс	0.00023923	рт	0.00060344	ру	0.00280652
рф	2.86E-05	рх	5.71E-05	рц	3.21E-05	рч	0.00016425	рш	0.00021781	рщ	6.43E-05	ры	0.00248873
рь	0.00054988	рэ	7.14E-06	рю	0.00013568	ря	0.00148896	са	0.00153894	сб	0.00013211	св	0.00142826
сг	0.00011426	сд	0.00048561	се	0.0034778	сж	8.21E-05	сз	6.07E-05	си	0.00178175	ск	0.00351708
сл	0.00427048	см	0.00122116	сн	0.00151395	со	0.0032707	сп	0.00271368	ср	0.00027494	сс	0.00068556
ст	0.0119545	су	0.00103191	сф	1.43E-05	сх	0.00016782	сц	2.86E-05	сч	0.00049989	сш	0.00011783
сщ	3.57E-06	сы	0.0005713	сь	0.00494176	сэ	3.21E-05	сю	0.00016782	ся	0.00510244	та	0.00647357
тб	0.00022852	тв	0.00269583	тг	0.00025709	тд	0.00060344	те	0.00492034	тж	0.00011783	тз	0.00022852
ти	0.0041955	тк	0.00152466	тл	0.00055702	тм	0.00055702	тн	0.00204955	то	0.01758539	тп	0.00088195
тр	0.00296006	тс	0.00309931	тт	0.00060701	ту	0.00171391	тф	3.57E-06	тх	0.00010712	тц	0.00014283
тч	0.00044276	тш	9.64E-05	тщ	2.14E-05	ты	0.00205312	ть	0.00883733	тэ	0.00016782	тю	4.64E-05
тя	0.00100692	та	0.00018924	уб	0.00096764	ув	0.0011069	уг	0.00128186	уд	0.00291007	уе	0.00033564
уж	0.00198527	уз	0.00049632	уй	0.00071413	уй	0.00023566	ук	0.00151752	ул	0.00189958	ум	0.00194243
ун	0.00134613	уо	0.00044633	уп	0.00166392	ур	0.00108904	ус	0.00211739	ут	0.00210311	уу	0.00015354
уф	1.43E-05	ух	0.00077483	уц	1.79E-05	уч	0.00135327	уш	0.00140683	ущ	0.0003285	уэ	8.21E-05
ую	0.00125686	уя	0.00024637	фа	6.43E-05	фе	2.14E-05	фи	9.64E-05	фл	1.43E-05	фо	0.00012854
фп	3.57E-06	фр	4.28E-05	фу	1.79E-05	фф	3.57E-06	фы	2.14E-05	ха	0.00061772	хб	0.00013211
хв	0.00050703	хг	0.00010712	хд	0.00022495	хе	0.00021067	хж	6.07E-05	хз	0.00010355	хи	0.00089266
хк	0.00028922	хл	0.00031065	хм	0.00032136	хн	0.000657	хо	0.00292792	хп	0.00027851	хр	0.00044276
хс	0.00033921	хт	0.0001464	ху	0.0002071	хф	1.43E-05	хх	3.93E-05	хц	3.21E-05	хч	8.93E-05
хш	4.64E-05	хщ	1.07E-05	хэ	1.79E-05	хю	1.43E-05	хя	4.64E-05	ца	0.0005963	цв	0.00028565
цг	1.07E-05	цд	7.14E-06	це	0.00081768	цз	1.43E-05	ци	0.00030707	цк	5.36E-05	цл	7.14E-06
цм	1.79E-05	цн	2.50E-05	цо	0.00036421	цп	2.50E-05	цр	1.07E-05	цс	3.57E-05	цт	3.21E-05

цу	0.00039634	цх	7.14E-06	цч	7.14E-06	цы	0.00021781	цэ	3.57E-06	ця	3.57E-06	ча	0.00191029
чв	1.79E-05	че	0.0035385	чж	3.57E-06	чз	1.07E-05	чи	0.0022138	чк	0.00058201	чл	4.28E-05
чн	0.00103548	чо	6.78E-05	чп	7.14E-06	чс	7.14E-06	чт	0.00420621	чу	0.00084624	чш	0.00014997
чь	0.00053917	чя	3.57E-06	ша	0.00125686	шб	1.07E-05	шв	2.86E-05	шд	1.07E-05	ше	0.00234591
шз	3.57E-06	ши	0.00233163	шк	0.00140683	шл	0.00060344	шм	3.21E-05	шн	0.00041776	шо	0.00042848
шп	2.50E-05	шр	2.86E-05	шс	7.14E-06	шт	3.57E-05	шу	0.00031422	шц	7.14E-06	шь	0.00078197
шэ	3.57E-06	шя	7.14E-06	ща	0.00045347	ще	0.00162821	щи	0.00112832	щк	3.57E-06	щм	3.57E-06
щн	5.71E-05	що	1.43E-05	щс	3.57E-06	щу	0.00023923	щь	3.21E-05	ья	0.0001464	ьб	0.000632
ыв	0.00159965	ыг	0.00028565	ыд	0.00031779	ые	0.00211739	ыж	0.00013568	ыз	0.00032136	ыи	0.00068199
ый	0.00199242	ык	0.00066414	ыл	0.00217452	ым	0.00170676	ын	0.00104977	ьо	0.00041776	ып	0.00078554
ыр	0.00051774	ыс	0.00111761	ыт	0.00109262	ыу	0.00017139	ыф	3.57E-06	ых	0.00144968	ыц	3.57E-06
ыч	0.0002678	ыш	0.00068199	ыщ	2.14E-05	ыз	5.36E-05	ью	3.57E-06	ья	0.00014283	ьа	0.00028922
ьб	0.00060701	ьв	0.00117831	ьг	0.00054631	ьд	0.00056773	ье	0.00128186	ьж	0.00018567	ьз	0.00064986
ьи	0.00146039	ьй	3.57E-06	ьк	0.00187458	ьл	0.00027851	ьм	0.00103906	ьн	0.00280652	ьо	0.0010819
ьп	0.00109976	ьр	0.00028922	ьс	0.00272083	ьт	0.00066057	ьу	0.00034635	ьф	2.50E-05	ьх	0.00012854
ьц	0.00027137	ьч	0.00089266	ьш	0.00056059	ьщ	1.79E-05	ьэ	0.00019996	ью	0.0005963	ья	0.00072127
эд	1.07E-05	эе	0.00028922	эй	3.93E-05	эк	2.50E-05	эм	2.50E-05	эн	2.50E-05	эо	3.57E-06
эп	1.43E-05	эр	1.43E-05	эс	2.50E-05	эт	0.00246731	эф	3.57E-06	эх	1.07E-05	эц	3.57E-06
юа	7.14E-05	юб	0.00045347	юв	0.00025352	юг	0.00010712	юд	0.00060344	юе	5.71E-05	юж	0.00013568
юз	9.28E-05	юи	0.00028208	юк	0.00033564	юл	6.07E-05	юм	0.00017853	юн	0.00041062	юо	0.00013925
юп	0.00028565	юр	0.00017853	юс	0.00073198	ют	0.00099621	юу	6.43E-05	юф	1.07E-05	юх	5.00E-05
юц	2.86E-05	юч	0.00032136	юш	5.00E-05	ющ	0.00022138	юз	3.21E-05	юю	2.14E-05	юя	0.00011426
яа	0.00026423	яб	0.00059273	яв	0.00189958	яг	0.0003892	яд	0.00147467	яе	0.00057844	яж	0.00043919
яз	0.00096407	яи	0.00115332	яй	6.07E-05	як	0.0009605	ял	0.00119259	ям	0.00093551	ян	0.00258157
яо	0.00101049	яп	0.00163178	яр	0.00055702	яс	0.00191386	ят	0.0023352	яу	0.0003892	яф	2.50E-05
ях	0.00031065	яц	7.14E-05	яч	0.000632	яш	0.00015711	ящ	0.00019639	яэ	0.00013925	яю	0.00016782
яя	0.00049275												

Код:

```
import re
from math import log

def import_data(filepath):
    with open(filepath, 'r', encoding='utf-8') as data_source:
        return data_source.read()

def calculate_freq_ngrams(text, ngram_len=1, step=1):

    freq = {}

    for i in range(0, len(text) - ngram_len + 1, step):
        ngram = text[i : i + ngram_len]
        if ngram in freq:
            freq[ngram] += 1
        else:
            freq[ngram] = 1

    for key in freq.keys():
        freq[key] = freq[key] * step / len(text)

    return freq

def calculate_entropy(freq_dict, ngram_len=1):
```

```

    entropy = 0
    for i in freq_dict.values():
        entropy += i * log(i, 2)

    return -entropy / ngram_len

def calculate_redundancy(entropy):
    return 1 - (entropy / log(33, 2))

def data_remove_spaces(text):
    return ''.join([letter for letter in text if letter != ' '])

def main():

    filepath = 'text_parsed.txt'
    data = import_data(filepath)

    # Text with spaces

    replaced_spaces_text = re.sub(' ', '_', data) # Replace spaces for better
    readability

    freq_letters = calculate_freq_ngrams(replaced_spaces_text)
    freq_bigrams = calculate_freq_ngrams(replaced_spaces_text, 2)
    freq_bigrams_step_2 = calculate_freq_ngrams(replaced_spaces_text, 2, 2)

    entropy_letters = calculate_entropy(freq_letters)
    entropy_bigrams = calculate_entropy(freq_bigrams, 2)
    entropy_bigrams_step_2 = calculate_entropy(freq_bigrams_step_2, 2)

    redundancy_letters = calculate_redundancy(entropy_letters)
    redundancy_bigrams = calculate_redundancy(entropy_bigrams)
    redundancy_bigrams_step_2 = calculate_redundancy(entropy_bigrams)

    # Text without spaces

    no_space_text = data_remove_spaces(data)

    freq_letters_nospace = calculate_freq_ngrams(no_space_text)
    freq_bigrams_nospace = calculate_freq_ngrams(no_space_text, 2)
    freq_bigrams_step_2_nospace = calculate_freq_ngrams(no_space_text, 2, 2)

    entropy_letters_nospace = calculate_entropy(freq_letters_nospace)
    entropy_bigrams_nospace = calculate_entropy(freq_bigrams_nospace, 2)
    entropy_bigrams_step_2_nospace = calculate_entropy(freq_bigrams_step_2_nospace,
2)

    redundancy_letters_nospace = calculate_redundancy(entropy_letters_nospace)
    redundancy_bigrams_nospace = calculate_redundancy(entropy_bigrams_nospace)
    redundancy_bigrams_step_2_nospace =
calculate_redundancy(entropy_bigrams_step_2_nospace)

main()

```

Висновок: виконавши комп'ютерний практикум, ми засвоїли поняття ентропії на символ джерела та його надлишковості, практично оцінили ці значення для довільного змістовного тексту (у різних моделях джерела).