

Міністерство освіти і науки України
Національний технічний університет України
“Київський політехнічний інститут ім. Ігоря Сікорського”
Фізико-технічний інститут

Лабораторна робота № 1
з предмету «Криптографія»
«Експериментальна оцінка ентропії на символ
джерела
відкритого тексту»

Виконали:
Студенти 3 курсу,
ФТІ, групи ФБ-74
Харламова Катерина, Стурчак Максим

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.

Постановка задачі:

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 $(10) H$, $(20) H$, $(30) H$.
3. Використовуючи отримані значення ентропії, оцінити надлишковість російської мови в різних моделях джерела.

	Без пробілів	З пробілами
Результат монограмми	4.4777346	4.389947
Надлишковість монограмми	0.11233451	0.12973753
Результат біграмми	4.167492	4.0044537
Надлишковість біграмми	0.17383698	0.20615765
Результат біграмми з накладанням	4.1740108	4.0099
Надлишковість біграмми з накладанням	0.17254467	0.20507796

Монограмма

Алфавит	Без пробелов	С пробелами
а	0,078743	0,065932
б	0,016092	0,013474
в	0,041650	0,034874
г	0,017944	0,015024
д	0,032278	0,027026
е	0,081415	0,068169
ж	0,010002	0,008375
з	0,015993	0,013391
и	0,065375	0,054738
й	0,011819	0,009896
к	0,034455	0,028849
л	0,053226	0,044566
м	0,033671	0,028193
н	0,064813	0,054268
о	0,111743	0,093563
п	0,028701	0,024031
р	0,044508	0,037267
с	0,051117	0,042800
т	0,062265	0,052135
у	0,029419	0,024633
ф	0,002225	0,001863
х	0,007929	0,006639
ц	0,003511	0,002940
ч	0,015593	0,013056
ш	0,008854	0,007413
щ	0,003373	0,002824
ъ	0,000298	0,000250
ы	0,017203	0,014404
ь	0,019868	0,016636
э	0,004700	0,003936
ю	0,006300	0,005275
я	0,024918	0,020864
пробел		0,162695

Біграмма без накладання

З пробілами			Без пробілів			
зс 0,000003	пн 0,000129	шт 0,000188	аа 0,000356	эф 0,000049	ою 0,000169	ця 0,000008
зт 0,000021	по 0,009622	шу 0,000243	аб 0,001700	зх 0,000008	па 0,001918	чб 0,000004
зу 0,000403	пп 0,000096	шц 0,000005	ав 0,005431	зц 0,000021	оя 0,001836	чв 0,000012
зч 0,000007	пр 0,006200	шч 0,000002	аг 0,001445	зч 0,000064	пб 0,000010	чд 0,000014
зш 0,000002	пс 0,000040	шь 0,000713	ад 0,003383	зш 0,000029	пв 0,000019	че 0,003832
зъ 0,000009	пт 0,000084	шэ 0,000002	ае 0,001681	зщ 0,000002	пд 0,000004	чз 0,000002
зы 0,000433	пу 0,000851	шю 0,000002	аж 0,001233	зъ 0,000010	пе 0,002297	чи 0,002221
зъ 0,000093	пч 0,000016	ща 0,000338	аз 0,003795	зы 0,0000517	пз 0,000010	чк 0,000622
зя 0,000188	пш 0,000016	ще 0,001351	аи 0,001247	зь 0,000111	пи 0,001350	чл 0,000021
иа 0,000200	пы 0,000379	щи 0,000851	ай 0,001074	зэ 0,000043	пк 0,000163	чм 0,000023
иб 0,000450	пъ 0,000083	щн 0,000021	ак 0,006083	зю 0,000002	пл 0,001192	чн 0,001327
ив 0,001880	пэ 0,000003	щю 0,000003	ал 0,011078	иа 0,000442	пм 0,000004	чо 0,000093
иг 0,000620	пю 0,000028	щр 0,000003	ам 0,004412	зя 0,000278	пн 0,000183	чп 0,000014
ид 0,001441	ра 0,007106	щу 0,000198	ан 0,006013	иб 0,001255	по 0,011488	чр 0,000012
ие 0,001682	ря 0,000545	щъ 0,000022	ао 0,001190	ив 0,004167	пп 0,000134	чс 0,000016
иж 0,000291	рб 0,000045	ье 0,000150	ап 0,003083	иг 0,001157	пр 0,007409	чт 0,003955
из 0,002035	рв 0,000424	ья 0,000088	ар 0,003818	ид 0,002830	пс 0,000051	чу 0,000739
ии 0,000551	рг 0,000190	яа 0,000002	ас 0,005911	ие 0,002435	пт 0,000107	чф 0,000002
ий 0,001142	рд 0,000400	ыб 0,000357	ат 0,006320	иж 0,000451	пу 0,001025	чх 0,000002
ик 0,001973	ре 0,005630	ыв 0,000906	ау 0,000661	из 0,002939	пч 0,000019	чш 0,000179
ил 0,005910	рж 0,000319	ыг 0,000195	аф 0,000354	ии 0,001663	пш 0,000019	чь 0,000317
им 0,002304	рз 0,000041	ыд 0,000105	ак 0,001239	ий 0,001369	пы 0,000451	ша 0,000920
ин 0,003943	ри 0,004751	ые 0,001010	ац 0,000300	ик 0,003488	пъ 0,000099	чя 0,000002
ио 0,000334	рк 0,000536	ыж 0,000033	ач 0,001406	ил 0,007339	пэ 0,000004	шб 0,000006
ип 0,000314	рл 0,000190	ыз 0,000105	аш 0,000990	им 0,003593	пю 0,000033	шв 0,000041
ир 0,000943	рм 0,000355	ьи 0,000052	ащ 0,000331	ин 0,006544	ра 0,008598	шг 0,000006
ис 0,002962	рн 0,001006	ый 0,001565	аз 0,000461	ио 0,001854	пя 0,000652	шд 0,000014
ит 0,003825	ро 0,007112	ык 0,000198	аю 0,001264	ип 0,002525	рб 0,000088	ше 0,002463
иу 0,000003	рп 0,000072	ыл 0,000694	ая 0,002422	ир 0,001585	рв 0,000609	шз 0,000002
ию 0,000053	рр 0,000476	ым 0,001170	ба 0,001115	ис 0,005225	рг 0,000263	ши 0,001622
ик 0,001148	рс 0,000377	ын 0,000121	бб 0,000082	ит 0,005452	рд 0,000560	шк 0,000794
ич 0,000743	рт 0,000905	ыо 0,000007	бв 0,000047	иу 0,000533	ре 0,006738	шл 0,000823
иш 0,001167	ру 0,002568	ып 0,000196	бг 0,000008	иф 0,000198	рж 0,000391	шм 0,000016
ищ 0,000679	рф 0,000014	ыр 0,000281	бд 0,000021	ик 0,001560	рз 0,000082	шн 0,000220
из 0,000198	рх 0,000105	ыс 0,000545	бе 0,002177	иц 0,000924	ри 0,005775	шо 0,000403
иэ 0,000014	рц 0,000021	ыт 0,000600	бж 0,000014	ич 0,001893	рк 0,000747	шп 0,000039
ию 0,000338	рч 0,000084	ыу 0,000007	бз 0,000023	иш 0,000971	рл 0,000253	шр 0,000021
ия 0,000996	рш 0,000219	ыф 0,000002	бн 0,001517	ищ 0,000247	рм 0,000457	шс 0,000012
иб 0,000016	рщ 0,000096	ых 0,000832	бк 0,000230	из 0,000358	рн 0,001272	шт 0,000224
йб 0,000024	ры 0,001442	ыц 0,000005	бл 0,001058	ию 0,000414	ро 0,008561	шу 0,000290
йв 0,000016	рь 0,000472	ыч 0,000129	бм 0,000019	ия 0,001902	рп 0,000216	шф 0,000002
йг 0,000002	рз 0,000246	ыш 0,000500	бн 0,000346	йа 0,000179	рр 0,000593	шц 0,000006
йд 0,000258	рю 0,000133	ыщ 0,000014	бо 0,002914	йб 0,000338	рс 0,000510	шч 0,000004
йе 0,000010	ря 0,001391	ыз 0,000003	бп 0,000019	йв 0,000634	рт 0,001144	шш 0,000004
йз 0,000048	сб 0,000803	ыя 0,000047	бр 0,001389	йг 0,000268	ру 0,003091	шь 0,000854
ий 0,000014	сб 0,000165	ыо 0,000003	бс 0,000214	йд 0,000854	рф 0,000035	шэ 0,000002
ий 0,000003	св 0,000960	ыб 0,000055	бт 0,000004	йе 0,000080	рх 0,000142	шо 0,000002
ий 0,000374	сг 0,000031	ыв 0,000095	бу 0,001313	йж 0,000144	рц 0,000027	шы 0,000002
ил 0,000060	сд 0,000248	ыг 0,000102	бф 0,000006	йз 0,000294	рч 0,000111	ша 0,000401
им 0,000109	се 0,003155	ыд 0,000383	бх 0,000072	ий 0,000447	рш 0,000276	шд 0,000002
ин 0,000343	сж 0,000019	ые 0,000362	бц 0,000004	ий 0,000004	рщ 0,000117	ше 0,001605
ио 0,000052	сз 0,000007	ыз 0,000136	бч 0,000006	ик 0,001115	ры 0,001725	ши 0,001021
ип 0,000014	си 0,001554	ыи 0,000053	бш 0,000010	ил 0,000218	рь 0,000564	шм 0,000002
ир 0,000014	ск 0,002816	ык 0,001013	бщ 0,000148	йм 0,000613	рз 0,000331	шн 0,000027
ис 0,000501	сл 0,003043	ыл 0,000003	бъ 0,000132	ин 0,000990	ро 0,000158	шп 0,000004
ит 0,000364	см 0,000858	ым 0,000298	бы 0,002527	ио 0,000506	ря 0,001008	шр 0,000004
иу 0,000003	сн 0,000870	ын 0,001384	бь 0,000072	ип 0,001039	св 0,001710	шс 0,000239
иф 0,000014	со 0,002835	ыо 0,000034	бэ 0,000117	ир 0,000403	сб 0,000315	шк 0,000002
ич 0,000007	сп 0,001397	ып 0,000029	бю 0,000023	ис 0,001457	св 0,001336	шц 0,000027
иц 0,000033	сс 0,000817	ыр 0,000005	ва 0,006559	ит 0,000757	сг 0,000144	ье 0,000179
ич 0,000095	ст 0,000958	ыс 0,001017	бя 0,000403	йу 0,000284	сд 0,000584	ья 0,000054
йш 0,000055	су 0,000639	ыт 0,000152	вб 0,000241	йф 0,000107	сж 0,000388	зя 0,000105
йз 0,000016	сф 0,000033	ыу 0,000005	вв 0,000414	йх 0,000072	ск 0,000084	ыб 0,000562
ка 0,006212	сх 0,000133	ыф 0,000017	вг 0,000436	йц 0,000088	сз 0,000082	ыв 0,001655
кв 0,000022	ц 0,000074	ыц 0,000138	вд 0,000862	йч 0,000387	си 0,002002	ыг 0,000340
кб 0,000005	ц 0,000262	ыч 0,000091	ве 0,005271	йш 0,000208	ск 0,000639	ыд 0,000391
кв 0,000196	сш 0,000095	ыш 0,000660	вж 0,000080	йщ 0,000012	сл 0,003702	ые 0,001346
кг 0,000003	сы 0,000045	ыщ 0,000005	вз 0,000591	йэ 0,000099	см 0,001206	ыж 0,000078
кд 0,000024	сь 0,000322	ыз 0,000005	ви 0,003920	ию 0,000006	сн 0,001463	ыз 0,000362
ке 0,001130	съ 0,002905	ыю 0,000608	вк 0,001097	ка 0,007503	со 0,003575	ыи 0,000436
кж 0,000047	сэ 0,000043	ья 0,000281	вл 0,000871	кб 0,000191	сп 0,002048	ый 0,001869
кз 0,000016	сю 0,000121	эб 0,000003	вм 0,000543	кб 0,000196	сп 0,000383	ык 0,000455
ки 0,002633	та 0,005852	эн 0,000010	вн 0,001766	кв 0,000617	сс 0,001251	ыл 0,000873
кл 0,000050	сб 0,004913	эг 0,000007	во 0,006361	кг 0,000105	ст 0,011759	ым 0,001671
кл 0,000896	тб 0,000162	эд 0,000060	вп 0,001017	кд 0,000342	су 0,000891	ын 0,000747
км 0,000047	тв 0,001752	эе 0,000002	вр 0,001117	ке 0,001426	сф 0,000080	ыо 0,000370
кн 0,000412	тг 0,000029	эз 0,000003	вс 0,002848	кж 0,000117	ск 0,000183	ып 0,000918
ко 0,008085	тд 0,000122	эй 0,000059	вт 0,001066	кз 0,000144	сц 0,000097	ыр 0,000416
кп 0,000002	те 0,005021	эк 0,000079	ву 0,000891	ки 0,003474	сч 0,000403	ыс 0,001183
кр 0,001780	тж 0,000002	эл 0,000698	вф 0,000086	кк 0,000253	сш 0,000134	ыт 0,000938
кс 0,000289	ти 0,003612	эм 0,000065	вх 0,000103	кл 0,001155	сь 0,000054	ыу 0,000189
ку 0,000636	тк 0,000836	эн 0,000415	вц 0,000066	км 0,000313	сы 0,000385	ыф 0,000012
кы 0,002164	тл 0,000212	еп 0,000053	вч 0,000187	кн 0,000871	сь 0,003468	ых 0,001089
кф 0,000002	тм 0,000090	эр 0,000076	вш 0,000887	ко 0,009977	сэ 0,000169	ыц 0,000010
кк 0,000002	тн 0,001134	эс 0,000022	вщ 0,000023	кп 0,000424	сю 0,000146	ыч 0,000253
кц 0,000033	то 0,013713	эт 0,002269	въ 0,000008	кр 0,002235	та 0,007090	ыш 0,000617
кч 0,000003	тп 0,000202	эф 0,000038	вы 0,003159	кс 0,000737	сб 0,005923	ыщ 0,000016
кш 0,000002	тр 0,002860	эх 0,000003	вь 0,000358	кт 0,001050	тб 0,000360	ыз 0,000126
кщ 0,000009	тс 0,001148	эш 0,000005	вэ 0,000259	ку 0,002702	тв 0,002793	ыю 0,000002
кь 0,000009	тту 0,000283	ээ 0,000002	вю 0,000006	кф 0,000070	тг 0,000179	ыз 0,000179
кэ 0,000017	ты 0,001463	юа 0,000005	га 0,001807	кх 0,000025	тд 0,000541	ыя 0,000198
кю 0,000009	тф 0,000043	яа 0,000005	вя 0,000496	кц 0,000056	те 0,006123	ыб 0,000391
ла 0,005450	тх 0,000009	юб 0,000246	гб 0,000101	кч 0,000185	тж 0,000148	ыв 0,001311
кя 0,000007	тц 0,000090	юв 0,000009	гв 0,000115	кш 0,000014	тз 0,000154	ыг 0,000290
лб 0,000047	тч 0,000124	юг 0,000028	гг 0,000103	кщ 0,000014	ти 0,004653	ыд 0,001058
лв 0,000022	тш 0,000009	юд 0,000364	гд 0,001581	кь 0,000010	тк 0,001301	ые 0,000846
лг 0,000133	тщ 0,000007	юе 0,000003	ге 0,000338	кз 0,000150	тл 0,000358	ыж 0,000097
лд 0,000078	ты 0,000010	юж 0,000031	гж 0,000006	кю 0,000021	тм 0,000420	ыз 0,000517
ле 0,004396	тэ 0,001728	юз 0,000009	гз 0,000035	ла 0,006596	тн 0,001937	ыи 0,000774
лж 0,000219	тэ 0,005569	юи 0,000034	гк 0,000768	кя 0,000208	то 0,016764	ык 0,001803
лз 0,000048	тэ 0,000005	юй 0,000029	гк 0,000119	лб 0,000366	тп 0,000846	ыл 0,000245
ли 0,006669	тя 0,000047	юк 0,000060	гл 0,001702	лв 0,000938	тр 0,003602	ым 0,000910
лм 0,000405	тю 0,000644	юл 0,000028	гм 0,000037	лг 0,000432	тс 0,001856	ын 0,002778
лн 0,000787	уа 0,000155	юм 0,000048	го 0,000331	лд 0,000510	тт 0,000584	ыо 0,000741
ло 0,000053	уб 0,000782	юн 0,000064	гп 0,008518	ле 0,005686	ту 0,001893	ып 0,001134
лп 0,000352	ув 0,000774	юп 0,000005	гп 0,000152	лж 0,000311	тф 0,000093	ыр 0,000220

ло 0,005487	үг 0,000763	юр 0,000078	гр 0,001126	лз 0,000298	тх 0,000060	ьс 0,002151
лп 0,000036	үд 0,001520	юс 0,000219	гс 0,000140	ли 0,008421	тц 0,000146	ьт 0,000673
лр 0,000009	үе 0,000133	ют 0,000348	гт 0,000043	лк 0,001099	тч 0,000356	ьу 0,000309
лс 0,002304	үж 0,001230	юх 0,000029	гү 0,000741	лл 0,001078	тш 0,000088	ьф 0,000076
лт 0,000183	үз 0,000238	юц 0,000007	гх 0,000012	лм 0,000634	тщ 0,000010	ьк 0,000109
лү 0,001334	үй 0,000176	юч 0,000210	гч 0,000054	лн 0,001408	ть 0,000012	ьц 0,000218
лф 0,000031	үй 0,000105	юш 0,000016	гш 0,000006	ло 0,007125	ты 0,002062	ьч 0,000568
лх 0,000003	үк 0,000813	ющ 0,000512	гэ 0,000016	лп 0,000914	ть 0,000656	ьш 0,000844
лч 0,000053	үл 0,002018	юэ 0,000009	гя 0,000037	лр 0,000327	тэ 0,000142	ьщ 0,000006
лы 0,000007	үт 0,001229	юю 0,000038	да 0,005520	лс 0,003443	тю 0,000060	ьз 0,000202
ль 0,000467	үн 0,000341	юя 0,000014	дб 0,000115	лт 0,000652	тя 0,000916	ью 0,000731
лэ 0,000026	үо 0,000093	яа 0,000005	дв 0,001389	лү 0,001786	үа 0,000303	ья 0,000652
лю 0,001024	үп 0,000817	яб 0,000119	дг 0,000064	лф 0,000103	үб 0,001107	эб 0,000006
ля 0,001673	үр 0,000560	яв 0,000272	дд 0,000138	лх 0,000060	үв 0,001601	эв 0,000012
мз 0,003022	үс 0,001353	яг 0,000074	де 0,006437	лц 0,000023	үг 0,001027	эг 0,000008
мб 0,000243	үт 0,001653	яд 0,000555	дж 0,000875	лч 0,000786	үд 0,002188	эд 0,000074
мв 0,000014	үү 0,000009	яе 0,000059	дз 0,000084	лш 0,000066	үе 0,000257	эе 0,000004
мг 0,000024	үф 0,000048	яж 0,000096	ди 0,002933	лщ 0,000006	үж 0,001527	эз 0,000004
мд 0,000017	үх 0,000443	яз 0,000214	дк 0,000342	лы 0,001048	үз 0,000449	эй 0,000070
ме 0,004175	үч 0,000002	яи 0,000016	дл 0,000885	лэ 0,000594	үй 0,000720	эк 0,000093
мж 0,000002	үш 0,001101	яй 0,000041	дм 0,000132	лз 0,000253	үй 0,000123	эл 0,000848
ми 0,003143	үш 0,000486	яи 0,000164	дн 0,002321	лю 0,001227	үк 0,001325	эм 0,000072
мл 0,000083	үщ 0,000312	ял 0,001084	до 0,004458	ля 0,002474	үл 0,002527	эн 0,000506
мк 0,000119	үэ 0,000041	ям 0,000302	дп 0,000346	ма 0,003748	үм 0,001930	эп 0,000070
мм 0,000098	үю 0,001173	ян 0,000665	др 0,001130	мб 0,000517	үн 0,001204	эр 0,000093
мн 0,001882	үя 0,000083	юо 0,000016	дс 0,000498	мв 0,000766	үо 0,000539	эс 0,000031
мо 0,004304	фа 0,000200	яп 0,000121	дт 0,000247	мг 0,000309	үп 0,001609	эт 0,002688
мп 0,000240	фе 0,000248	яр 0,000102	дү 0,001988	мд 0,000418	үр 0,000827	эф 0,000045
мр 0,000040	фи 0,000264	яс 0,000424	дф 0,000012	ме 0,005091	үс 0,002190	эх 0,000004
мс 0,000150	фк 0,000007	ят 0,001851	дх 0,000086	мж 0,000084	үт 0,002276	эш 0,000006
мт 0,000078	фл 0,000060	юу 0,000002	дц 0,000440	мз 0,000189	үү 0,000154	эз 0,000002
му 0,001792	фм 0,000002	як 0,000140	дч 0,000054	ми 0,004252	үф 0,000095	юз 0,000041
мф 0,000007	фн 0,000007	яц 0,000071	дш 0,000082	мк 0,000562	үх 0,000576	эя 0,000006
мх 0,000002	фо 0,000383	яч 0,000159	дъ 0,000076	мл 0,000272	үц 0,000016	юб 0,000368
мц 0,000019	фп 0,000005	яш 0,000017	ды 0,000681	мм 0,000508	үч 0,001906	юв 0,000265
мш 0,000024	фр 0,000381	ящ 0,000160	дь 0,000426	мн 0,003075	үш 0,000628	юг 0,000099
мч 0,000005	фт 0,000029	яз 0,000007	дэ 0,000043	мо 0,005701	үщ 0,000379	юд 0,000626
мщ 0,000002	фү 0,000140	яю 0,000090	дю 0,000058	мп 0,001208	үэ 0,000185	юе 0,000070
мы 0,000748	фф 0,000053	яя 0,000115	еа 0,000377	мр 0,000294	үйю 0,001401	юж 0,000113
мь 0,000212	фч 0,000003	Пробел 0,003653	дя 0,000354	мс 0,000992	үя 0,000286	юз 0,000068
мэ 0,000138	фы 0,000017		еб 0,001998	мт 0,000393	фа 0,000237	юи 0,000228
мю 0,000002	фх 0,000009		ев 0,003704	мү 0,002402	фд 0,000004	юй 0,000035
на 0,009767	фа 0,000579		ег 0,003499	мф 0,000056	фе 0,000298	юк 0,000331
ня 0,000453	хб 0,000002		ед 0,003933	мх 0,000074	фи 0,000317	юл 0,000093
нв 0,000017	хв 0,000264		ее 0,001725	мц 0,000037	фк 0,000014	юм 0,000214
нг 0,000062	хг 0,000003		еж 0,000996	мч 0,000327	фл 0,000068	юн 0,000364
нз 0,000593	хд 0,000003		ез 0,002227	мш 0,000058	фм 0,000002	юо 0,000208
нд 0,000420	хе 0,000053		ей 0,001037	мщ 0,000016	фн 0,000010	юп 0,000321
не 0,009076	жж 0,000002		ей 0,002282	мы 0,000897	фо 0,000453	юр 0,000189
нж 0,000034	хз 0,000002		ек 0,002468	мь 0,000253	фп 0,000008	юс 0,000566
нз 0,000053	хи 0,000186		ел 0,006824	мэ 0,000272	фр 0,000455	ют 0,000541
ни 0,006827	хл 0,000005		ем 0,005147	мю 0,000010	фс 0,000002	юу 0,000058
нк 0,000589	хм 0,000150		ен 0,010601	на 0,011694	фт 0,000037	юф 0,000031
нл 0,000076	хм 0,000045		ео 0,001474	мя 0,000827	фү 0,000163	юх 0,000064
нм 0,000002	хн 0,000233		еп 0,003215	нб 0,000144	фф 0,000064	юц 0,000025
нн 0,002595	хо 0,002066		ер 0,008572	нв 0,000438	фч 0,000004	юч 0,000428
но 0,009422	хп 0,000005		ес 0,007102	нг 0,000766	фы 0,000021	юш 0,000066
нп 0,000005	хр 0,000169		ет 0,006903	нд 0,000648	фь 0,000010	ющ 0,000613
нр 0,000134	хс 0,000045		еу 0,000729	не 0,010895	ха 0,000733	юз 0,000062
нс 0,000601	хт 0,000024		еф 0,000181	нж 0,000070	фя 0,000002	юю 0,000049
нт 0,000968	ху 0,000129		ек 0,000998	нз 0,000206	хб 0,000115	юя 0,000115
нү 0,002790	хх 0,000002		ец 0,000354	ни 0,008401	хв 0,000578	яз 0,000183
нф 0,000081	хч 0,000007		еч 0,001797	нк 0,000838	хг 0,000088	яб 0,000619
нх 0,000047	хш 0,000014		еш 0,001229	нл 0,000134	хд 0,000167	яв 0,002056
нц 0,000286	хъ 0,000005		ещ 0,000772	нм 0,000158	хе 0,000103	яг 0,000331
нч 0,000240	хь 0,000009		ез 0,000288	нн 0,003488	жж 0,000033	яд 0,001531
нш 0,000016	хэ 0,000065		ею 0,000136	но 0,011490	хз 0,000093	яе 0,000383
нщ 0,000128	хя 0,000012		жа 0,001517	нп 0,000576	хи 0,000385	яж 0,000282
ны 0,002929	ца 0,000657		яб 0,000784	нр 0,000245	жк 0,000224	яз 0,000943
нэ 0,001249	цб 0,000096		жб 0,000025	нс 0,001031	хл 0,000268	яи 0,000838
ня 0,000005	цг 0,000003		жв 0,000033	нт 0,001288	жм 0,000204	ий 0,000049
ню 0,000209	цд 0,000003		жг 0,000010	нү 0,003492	жн 0,000506	як 0,001006
нй 0,001894	це 0,000898		жд 0,000770	нф 0,000111	хо 0,002628	ял 0,001469
об 0,000017	цз 0,000002		же 0,003484	нх 0,000113	хп 0,000327	ям 0,001017
оа 0,003472	ца 0,000505		жж 0,000004	нц 0,000356	хр 0,000280	ян 0,002715
ов 0,006112	цк 0,000014		жз 0,000012	нч 0,000356	хс 0,000319	яо 0,000998
ог 0,004758	цл 0,000002		жи 0,001801	нш 0,000039	хт 0,000150	яп 0,002328
од 0,005101	цп 0,000171		жк 0,000228	нщ 0,000154	ху 0,000243	яр 0,000508
ое 0,001804	цр 0,000002		жл 0,000023	ны 0,003499	хф 0,000023	яс 0,001910
ож 0,002026	цс 0,000005		жм 0,000025	нь 0,001490	хх 0,000041	ят 0,002854
оз 0,001268	цт 0,000009		жн 0,001037	нэ 0,000058	хц 0,000012	яү 0,000624
ой 0,000619	цу 0,000133		жо 0,000366	ню 0,000251	хч 0,000084	яф 0,000051
ои 0,004130	чу 0,000002		жп 0,000037	ня 0,002319	шш 0,000051	ях 0,000354
ок 0,002164	цы 0,000177		жр 0,000008	оз 0,000220	хъ 0,000006	яц 0,000109
ол 0,006012	ча 0,001775		жс 0,000039	об 0,004875	хь 0,000010	яч 0,000640
ом 0,006133	чв 0,000002		жт 0,000025	ов 0,009473	хз 0,000111	яш 0,000084
он 0,005605	чб 0,003209		жү 0,000276	ог 0,006164	хя 0,000093	ящ 0,000196
оо 0,000174	чи 0,001839		жч 0,000158	од 0,007129	ца 0,000796	яз 0,000286
оп 0,001272	чк 0,000515		жъ 0,000037	ое 0,002616	цб 0,000006	яю 0,000113
ор 0,005276	чл 0,000017		жэ 0,000012	ож 0,002710	цв 0,000136	яя 0,000377
ос 0,006047	чм 0,000012		за 0,005618	оз 0,002243	цг 0,000012	
от 0,007261	чн 0,001105		жя 0,000025	ои 0,001657	цд 0,000019	
оу 0,000214	чо 0,000067		зб 0,000158	ой 0,004929	це 0,001072	
оф 0,000124	чр 0,000009		зв 0,001004	ок 0,003636	цз 0,000010	
ох 0,000553	чт 0,003326		зг 0,000379	ол 0,007569	ци 0,000638	
оц 0,000103	чү 0,000619		зд 0,001066	ом 0,008633	цк 0,000025	
оч 0,001956	чш 0,000150		зе 0,000393	он 0,008796	цл 0,000002	
ош 0,000941	чь 0,000265		зж 0,000142	оо 0,001712	цм 0,000012	
ощ 0,000238	ша 0,000769		зз 0,000111	оп 0,003830	цн 0,000023	
оз 0,000088	ча 0,000002		зи 0,000714	ор 0,006880	цо 0,000226	
ою 0,000134	шв 0,000029		зк 0,000315	ос 0,009168	цп 0,000029	
ла 0,001596	ше 0,002056		зл 0,000202	от 0,009864	цр 0,000016	
ля 0,000819	ши 0,001341		зм 0,000689	оу 0,000920	цс 0,000021	
лб 0,000007	шк 0,000667		зн 0,001856	оф 0,000276	цт 0,000019	
лв 0,000003	шл 0,000686		зо 0,000805	ох 0,000866	цу 0,000173	
ле 0,001920	шм 0,000007		зп 0,000222	оц 0,000173	цф 0,000004	
ли 0,001124	шн 0,000172		зр 0,000270	оч 0,003141	цш 0,000002	
лп 0,000122	шо 0,000333		зс 0,000163	ош 0,001257	цы 0,000212	
лр 0,000998	шт 0,000016		зт 0,000117	ощ 0,000305	цэ 0,000002	
лм 0,000002	шр 0,000014		зү 0,000574	оз 0,000634	ча 0,002128	

Біграмма з накладанням

3 проблеми						Без проблів									
аа 0,000095	вв 0,000300	ол 0,006014	уо 0,000093	юз 0,000009	аа 0,000358	жя 0,000025	от 0,009874	хш 0,000051	яв 0,002062						
аб 0,000987	аз 0,000002	ом 0,006136	уо 0,000819	юи 0,000034	аб 0,001704	зб 0,000158	оу 0,009020	хш 0,000006	яв 0,000331						
ав 0,003369	во 0,000002	он 0,005614	уо 0,000562	юй 0,000031	ав 0,005437	зв 0,001006	оф 0,000276	хш 0,000010	яд 0,001533						
аг 0,000855	ва 0,000333	оо 0,000174	уо 0,001354	юк 0,000060	аг 0,001445	зг 0,000379	ох 0,000866	хш 0,000113	яв 0,000383						
ад 0,002028	га 0,001504	оп 0,001273	уо 0,001654	юл 0,000028	ад 0,003383	зд 0,001366	оц 0,000173	хя 0,000093	яж 0,000282						
ае 0,001074	гб 0,000009	ос 0,005283	уу 0,000009	юм 0,000048	ае 0,001686	зе 0,000397	оч 0,003149	ца 0,000803	зв 0,000945						
аж 0,000941	гв 0,000012	ос 0,006048	уф 0,000048	юн 0,000064	аж 0,001233	жж 0,000142	ош 0,001257	цб 0,000006	яю 0,000840						
ак 0,002766	гг 0,000081	от 0,002761	уф 0,000443	юп 0,000005	зз 0,003795	зг 0,000111	оц 0,000205	цв 0,000142	яй 0,000049						
ай 0,000091	гд 0,001273	оу 0,000214	уи 0,000002	юр 0,000083	ий 0,001247	зк 0,000718	оо 0,000634	цг 0,000012	як 0,001010						
аи 0,000898	ге 0,000264	оф 0,000124	уи 0,001103	юс 0,000222	ай 0,001074	зк 0,000315	ою 0,000169	цд 0,000019	ял 0,001471						
ал 0,004151	гк 0,000610	ох 0,000555	уш 0,000488	ют 0,000350	ак 0,006086	зл 0,000204	па 0,001920	це 0,001093	ям 0,001023						
ам 0,009033	гк 0,000078	оц 0,000103	уш 0,000314	юу 0,000029	ал 0,011084	зм 0,000689	он 0,001836	цж 0,000010	яд 0,002727						
ан 0,002795	гл 0,001430	оч 0,001956	уу 0,000041	юц 0,000007	ам 0,004412	зн 0,001860	пб 0,000010	цк 0,000646	яю 0,001000						
ао 0,003631	гм 0,000002	ош 0,000941	ую 0,001177	юч 0,000210	ао 0,006024	зо 0,000807	пв 0,000019	цм 0,000025	яв 0,002342						
ап 0,000067	гн 0,000202	оц 0,000238	фа 0,000205	юш 0,000016	ао 0,001201	зп 0,000222	пд 0,000004	цл 0,000002	яр 0,000508						
ап 0,000955	го 0,007091	оэ 0,000088	уя 0,000083	юу 0,000512	ап 0,003985	зр 0,000270	пе 0,002305	цм 0,000012	яс 0,001916						
ар 0,002805	гп 0,000003	ою 0,000134	фе 0,000257	юз 0,000009	ар 0,003820	зс 0,000163	пз 0,000010	цн 0,000023	ят 0,002857						
ас 0,003582	гр 0,009925	па 0,001599	фг 0,000267	юю 0,000026	ас 0,005915	зт 0,000117	пк 0,001354	цч 0,000026	яв 0,000624						
ат 0,004453	гс 0,000024	оя 0,000819	фк 0,000009	юя 0,000014	ат 0,006330	зу 0,000574	пл 0,000163	цп 0,000029	яф 0,000051						
ау 0,000136	гт 0,000016	пб 0,000007	фл 0,000062	яа 0,000002	ау 0,006661	зф 0,000049	пм 0,001194	цр 0,000016	яв 0,000354						
аф 0,000174	гу 0,000577	пв 0,000003	фм 0,000002	яб 0,000119	аф 0,000354	зх 0,000008	пн 0,000004	цс 0,000021	яй 0,000109						
ав 0,000886	гв 0,000033	пг 0,001927	фн 0,000007	яв 0,000274	ав 0,001241	зц 0,000021	по 0,000183	цт 0,000019	яч 0,000640						
ач 0,000191	гв 0,000003	пк 0,001125	фп 0,000388	яг 0,000074	ач 0,000300	зч 0,000064	пй 0,001152	цу 0,000173	яш 0,000084						
ац 0,000817	гз 0,000003	пл 0,000124	фп 0,000027	яд 0,000555	ац 0,001406	зо 0,000029	пн 0,000136	цф 0,000004	ящ 0,000019						
аш 0,000679	га 0,004606	пл 0,000999	фр 0,000386	яе 0,000059	аш 0,000992	зиц 0,000002	пс 0,000743	цш 0,000002	яю 0,000286						
аш 0,000269	гя 0,000007	пм 0,000002	фр 0,000029	жж 0,000096	аш 0,000333	зо 0,000010	пс 0,000054	цш 0,000016	яю 0,000113						
зб 0,000017	дб 0,000062	пн 0,000129	фу 0,000140	зз 0,000214	зб 0,000461	зю 0,000517	пт 0,000107	цш 0,000004							
зо 0,001037	де 0,001087	по 0,009634	фб 0,000055	ян 0,000126	зо 0,001126	зю 0,000111	пу 0,001025	ца 0,002136							
ав 0,001527	дг 0,000031	пп 0,000098	фг 0,000003	яй 0,000041	ад 0,002424	зз 0,000043	пн 0,000019	цж 0,000008							
ба 0,000925	да 0,000052	пр 0,006212	фы 0,000017	як 0,000165	ба 0,001117	зю 0,000002	пш 0,000019	цб 0,000004							
бб 0,000067	де 0,005385	пс 0,000041	фы 0,000009	ял 0,001086	бб 0,000082	ио 0,000445	пш 0,000453	цв 0,000012							
бв 0,000033	дж 0,000729	пт 0,000084	ха 0,000584	ям 0,000305	бв 0,000047	зя 0,000280	пш 0,000099	цд 0,000019							
бг 0,000003	дз 0,000007	пу 0,000856	хб 0,000002	яп 0,000665	бг 0,000008	ио 0,001257	пш 0,000004	цж 0,003844							
ба 0,000014	ди 0,002406	пч 0,000016	хв 0,000269	яо 0,000016	ба 0,000021	ив 0,004172	пю 0,000033	цз 0,000002							
бв 0,001830	дк 0,0000221	пш 0,000016	хг 0,000003	яп 0,000122	бв 0,002192	ий 0,001159	ра 0,008617	цк 0,002231							
бж 0,000012	дл 0,000707	пш 0,000379	хд 0,000003	яр 0,000102	бж 0,000014	ид 0,002830	пш 0,000652	цк 0,000622							
бз 0,000016	дм 0,0000050	пш 0,000005	хе 0,000055	яс 0,000427	бз 0,000023	ие 0,002437	рб 0,000888	чл 0,000021							
бн 0,001263	дн 0,001827	пз 0,000003	хж 0,000002	яч 0,001854	бн 0,001521	иж 0,000451	рв 0,000609	чм 0,000023							
бк 0,000191	до 0,003703	пз 0,000028	хз 0,000002	яч 0,000002	бк 0,000233	из 0,002947	рз 0,000263	чн 0,001327							
бл 0,000884	дп 0,000193	ра 0,007124	хи 0,000186	як 0,000140	бл 0,001058	ий 0,001631	рб 0,000562	чо 0,000093							
бм 0,000012	др 0,000917	ля 0,000545	хл 0,000005	яц 0,000071	бм 0,000019	ид 0,001372	рж 0,000372	чс 0,000014							
бн 0,000276	дс 0,000317	рб 0,000045	хм 0,000150	яо 0,000159	бн 0,000346	ик 0,003488	рж 0,000391	чр 0,000012							
бо 0,002418	дт 0,001513	рв 0,000424	хм 0,000045	яш 0,000159	бо 0,002931	ил 0,007345	роз 0,000348	чс 0,000016							
бр 0,001167	дү 0,001628	рг 0,000190	хн 0,000233	ящ 0,000160	бр 0,000019	им 0,003595	рп 0,005779	чт 0,003978							
бс 0,000171	дф 0,000002	рд 0,000401	хо 0,000278	ящ 0,000007	бр 0,001395	рп 0,006571	рк 0,000747	чы 0,000741							
бү 0,001091	дх 0,000060	ре 0,005638	хп 0,000005	ю 0,000090	бс 0,000216	ио 0,001856	рл 0,000255	чф 0,000002							
бф 0,000002	дц 0,000362	рж 0,000319	хр 0,000171	яя 0,000115	бр 0,000004	ип 0,002527	рм 0,000457	чс 0,000002							
бх 0,000057	дш 0,000017	рз 0,000041	хс 0,000045	Пробел	бү 0,001319	рп 0,001585	рн 0,001274	чш 0,000179							
бб 0,000002	дш 0,000062	ри 0,004759	хт 0,000026	0,003657	бф 0,000006	ис 0,005229	ро 0,008656	чс 0,000317							
бш 0,000007	дъ 0,000064	рл 0,000536	ху 0,000129		бх 0,000072	ит 0,005458	рп 0,000216	шз 0,000926							
бщ 0,000124	ды 0,000050	рл 0,000191	хл 0,000003		бц 0,000004	ир 0,000533	рр 0,000533	чс 0,000002							
бъ 0,000110	дъ 0,000357	рм 0,000355	хл 0,000007		бч 0,000006	иф 0,000198	рс 0,000510	шб 0,000006							
бы 0,002121	дз 0,000014	рн 0,001008	хш 0,000014		бш 0,000010	их 0,001564	рх 0,001564	шв 0,000043							
бв 0,000060	дю 0,000048	ро 0,007115	хы 0,000005		бщ 0,000148	иц 0,000926	рү 0,003101	шр 0,000006							
ба 0,000012	еа 0,000131	рп 0,000072	хп 0,000009		бщ 0,000132	ин 0,001897	рф 0,000035	шд 0,000014							
бж 0,000019	ед 0,000267	пт 0,000084	хр 0,000069		бы 40,002533	иш 0,000971	рх 0,000142	ше 0,002480							
ва 0,005387	еб 0,001063	пу 0,000856	хш 0,000012		бь 0,000072	ищ 0,000247	иц 0,000247	шз 0,000002							
ба 0,000331	ев 0,001491	пч 0,000016	ца 0,000663		бэ 0,000117	из 0,000358	рп 0,000111	ши 0,001630							
вб 0,000016	ег 0,002549	пш 0,000016	цв 0,000098		бю 0,000023	ию 0,000414	рш 0,000276	шж 0,000807							
вв 0,000078	ед 0,002504	пш 0,000379	цг 0,000003		ва 0,006563	ия 0,001902	рщ 0,000117	шл 0,000825							
вг 0,000021	ее 0,001161	пъ 0,000083	цд 0,000003		вб 0,000403	ия 0,000177	рщ 0,001727	шм 0,000019							
вд 0,000190	еж 0,000732	пз 0,000003	цз 0,000910		вб 0,000241	йб 0,000340	ръ 0,000564	шн 0,000220							
вз 0,004336	ез 0,001215	пю 0,000028	цз 0,000002		вв 0,000414	йв 0,000634	рз 0,000333	шо 0,000405							
вж 0,000003	ен 0,000053	ра 0,007124	цз 0,000512		вг 0,000438	йг 0,000268	ро 0,000158	шп 0,000039							
вй 0,000370	ей 0,001911	рл 0,000545	цч 0,000014		вд 0,000862	йд 0,000856	са 0,001716	шр 0,000021							
ви 0,0003078	ек 0,001399	рб 0,000045	цч 0,000002		ве 0,005275	йе 0,000084	рч 0,001010	шс 0,000012							
вк 0,000422	ел 0,005468	рв 0,000424	цш 0,000171		вж 0,000080	йж 0,000144	рф 0,000317	шт 0,000226							
вл 0,000582	ем 0,003295	рг 0,000190	цр 0,000002		вз 0,000591	йз 0,000294	св 0,001338	шу 0,000294							
во 0,000112	ен 0,007375	рд 0,000401	цр 0,000005		вз 0,003927	йз 0,000447	сг 0,000144	шф 0,000002							
вп 0,001065	ео 0,000162	ре 0,005638	цр 0,000009		вк 0,001097	йз 0,000004	сд 0,000584	шц 0,000006							
во 0,004977	еп 0,000719	рж 0,000319	цш 0,000134		вл 0,000873	йи 0,001115	се 0,003840	шч 0,000004							
вр 0,000210	ер 0,006712	рз 0,000041	цш 0,000002		вм 0,000545	йи 0,000220	сж 0,000086	шш 0,000004							
вп 0,000738	ес 0,004332	ри 0,004759	цш 0,000181		вн 0,001770	йм 0,000617	сз 0,000082	шы 0,000854							
вс 0,001816	ет 0,005002	рк 0,000536	ча 0,001789		во 0,006374	йо 0,000992	сн 0,002004	шю 0,000002							
вт 0,0000515	еу 0,000055	рл 0,000191	цш 0,000002		вп 0,001019	йю 0,000510	сн 0,003643	шю 0,000002							
вү 0,000646	еф 0,000086	рм 0,000355	че 0,003217		вр 0,001122	йп 0,001039	ск 0,003706	шз 0,000405							
вх 0,000028	ех 0,000589	рн 0,001008	чин 0,001854		вс 0,002857	йр 0,000403	см 0,001206	шн 0,000002							
вц 0,000024	ец 0,000252	рл 0,0007115	чк 0,000517		вт 0,001066	йс 0,001457	сн 0,001463	шд 0,000002							
вч 0,000048	еч 0,000941	рп 0,000072	чл 0,000017		вү 0,000891	йг 0,000757	со 0,003581	ше 0,001622							
вш 0,000636	еш 0,000903	рр 0,000476	чм 0,000012		вх 0,000086	йи 0,000284	ср 0,002052	ши 0,001037							
вщ 0,000005	еш 0,000632	рс 0,000377	чин 0,001106		вх 0,000103	йф 0,000107	ср 0,000385	шм 0,000002							
въ 0,000007	ез 0,000007	рт 0,000905	чл 0,000067		вц 0,000066	йц 0,000072	ср 0,001253	шн 0,000027							
вы 0,002647	ею 0,000112	рү 0,002576	чр 0,000009		вч 0,000187	йц 0,000088	сг 0,011774	шр 0,000004							
зб 0,000112	жа 0,001225														

из 0,002035	зб 0,000088	ск 0,000019	ща 0,000339	гр 0,001132	кт 0,001052	тм 0,000420	ыс 0,001185	
ии 0,000551	зв 0,000751	ж 0,001499	ще 0,001358	гс 0,000142	ку 0,002704	тн 0,001937	ыт 0,000938	
ий 0,001144	зг 0,000277	сз 0,000007	щи 0,000868	гт 0,000043	кф 0,000070	то 0,016783	ыу 0,000189	
ик 0,001973	зд 0,000772	з 0,004725	щн 0,000021	гү 0,000741	кх 0,000025	тп 0,000846	ыф 0,000012	
ил 0,005919	зе 0,000308	си 0,001556	що 0,000003	гх 0,000012	кц 0,000056	тр 0,003606	ых 0,001093	
им 0,002306	жк 0,000102	и 0,008282	щр 0,000003	гч 0,000054	кч 0,000185	тс 0,001858	ыц 0,000010	
ин 0,003948	зи 0,000064	ск 0,000284	щү 0,000202	гш 0,000006	кш 0,000014	тт 0,000584	ыч 0,000255	
ио 0,000338	зэ 0,000550	й 0,000005	щ 0,000022	гз 0,000016	кщ 0,000014	тү 0,001895	ыш 0,000617	
ип 0,000314	лп 0,000038	к 0,008587	ье 0,000160	да 0,005530	кь 0,000010	тф 0,000093	ыщ 0,000016	
ир 0,000944	лр 0,000009	сн 0,003045	ьн 0,000090	гя 0,000037	кэ 0,000150	тх 0,000060	ыз 0,000126	
ис 0,002964	лс 0,002304	см 0,000860	ьз 0,000002	дб 0,000115	кю 0,000021	тч 0,000146	ыю 0,000002	
ит 0,003829	лф 0,000183	л 0,002502	ьб 0,000357	дв 0,001391	ла 0,006606	тч 0,000356	ьа 0,000179	
иу 0,000003	лү 0,001337	м 0,008209	ьв 0,000908	дг 0,000064	кя 0,000208	тш 0,000088	ья 0,000198	
иф 0,000053	лф 0,000031	сн 0,000874	ьг 0,000195	дд 0,000138	лб 0,000366	тщ 0,000010	ьб 0,000391	
ик 0,001149	лх 0,000003	н 0,014768	ьд 0,000105	де 0,006446	лв 0,000938	ть 0,000012	ьв 0,001311	
иц 0,000743	лч 0,000053	со 0,000286	ьй 0,001010	дж 0,000877	лг 0,000432	ты 0,002064	ьг 0,000292	
ич 0,001172	лш 0,000007	о 0,009834	ьж 0,000033	дз 0,000084	лд 0,000510	ть 0,006660	ьд 0,001058	
иш 0,000679	лш 0,000877	сп 0,001397	ьз 0,000105	ди 0,002937	ле 0,005694	тэ 0,000142	ье 0,000850	
ищ 0,000200	лш 0,004268	п 0,017030	ьн 0,000052	дк 0,000342	лж 0,000311	тю 0,000060	ьж 0,000097	
из 0,000014	лэ 0,000026	ср 0,000212	ьй 0,001566	дл 0,000893	лз 0,000298	тя 0,000916	из 0,000519	
ию 0,000338	лю 0,001027	р 0,004158	ьн 0,000198	дм 0,000132	ли 0,008440	уа 0,000305	ьн 0,000774	
ия 0,000996	ля 0,001675	сс 0,000817	ьн 0,000696	дн 0,002321	лк 0,001099	уб 0,001109	ьк 0,001805	
йа 0,000016	мб 0,003035	ст 0,009669	ьн 0,001170	до 0,004472	лл 0,001080	ув 0,001603	ьл 0,000249	
йб 0,000024	мб 0,000243	с 0,013649	ьн 0,000121	дп 0,000346	лм 0,000636	уг 0,001027	ьм 0,000910	
йв 0,000016	мв 0,000014	т 0,007237	ьо 0,000007	др 0,001138	лн 0,001408	уд 0,002190	ьн 0,002780	
йг 0,000002	мг 0,000024	сү 0,000641	ьп 0,000196	дс 0,000500	ло 0,007133	уе 0,000261	ьо 0,000741	
йд 0,000262	ме 0,004182	ү 0,004520	ьр 0,000283	дт 0,000249	лп 0,000914	уж 0,001531	ьп 0,001136	
йе 0,000010	мк 0,000002	ф 0,000033	ьс 0,000546	дү 0,001990	лр 0,000327	юз 0,000451	ьр 0,000222	
йз 0,000048	мн 0,003147	ф 0,001037	ьт 0,000600	дф 0,000012	лс 0,000345	юз 0,000720	ьс 0,002153	
йн 0,000014	мн 0,000083	сх 0,000133	ьу 0,000007	дх 0,000086	лт 0,000652	ый 0,000126	ьт 0,000679	
йй 0,000003	мл 0,000121	ч 0,001558	ьф 0,000002	дц 0,000440	лү 0,001786	ын 0,001327	ьу 0,000309	
йк 0,000374	мм 0,000100	ц 0,000074	ьх 0,000834	дч 0,000834	лф 0,000103	ул 0,002527	ьф 0,000076	
йл 0,000062	мн 0,001882	сч 0,000262	ьц 0,000005	дш 0,000082	лх 0,000060	ум 0,001943	ьх 0,000109	
йм 0,000110	мо 0,004317	ц 0,000481	ьч 0,000129	дь 0,000076	лц 0,000023	ун 0,001208	ьц 0,000218	
йн 0,000345	мп 0,000240	мт 0,000532	ьш 0,000500	ды 0,000681	лч 0,000786	уо 0,000539	ьч 0,000568	
йо 0,000055	мр 0,000040	су 0,000095	ьщ 0,000014	дб 0,000426	лш 0,000066	уп 0,001611	ьш 0,000844	
ип 0,000014	мс 0,000152	ш 0,000129	ьз 0,000003	дз 0,000043	лщ 0,000006	ур 0,000827	ьщ 0,000006	
ир 0,000014	мт 0,000079	щ 0,000110	ья 0,000047	дю 0,000058	лы 0,001048	ус 0,002194	ьз 0,000202	
ис 0,000501	мү 0,001796	сб 0,000045	ья 0,000003	еа 0,000377	ль 0,000509	ут 0,002280	ью 0,000739	
ит 0,000364	мф 0,000007	сб 0,000322	ьб 0,000055	дя 0,000354	лэ 0,000253	үү 0,000154	ья 0,000652	
иу 0,000003	мх 0,000002	сб 0,0002905	ьв 0,000095	еб 0,001998	лю 0,001233	уф 0,000095	ьб 0,000006	
йф 0,000014	мц 0,000019	сз 0,000043	ьг 0,000102	ев 0,003704	ля 0,002474	ух 0,000576	ьв 0,000012	
йх 0,000007	мч 0,000024	з 0,003138	ьд 0,000384	ег 0,003505	ма 0,003760	уц 0,000016	ьг 0,000008	
йц 0,000033	мш 0,000005	сн 0,000121	ье 0,000365	ед 0,003933	мб 0,000517	уч 0,001908	ьд 0,000074	
ич 0,000095	мщ 0,000002	сн 0,004923	ьз 0,000136	ее 0,001725	мв 0,000766	уш 0,000632	ье 0,000004	
йш 0,000055	мы 0,000751	та 0,005857	ьн 0,000053	еж 0,000996	мг 0,000309	ущ 0,000381	ьз 0,000004	
йэ 0,000016	мь 0,000212	ю 0,000088	ьк 0,001015	ез 0,002229	мд 0,000418	юз 0,000185	ьй 0,000070	
ка 0,006224	мз 0,000138	я 0,004522	ьл 0,000003	ей 0,001037	ме 0,005104	ую 0,001406	ьж 0,000095	
ия 0,000022	мю 0,000002	тб 0,000162	ьм 0,000298	ей 0,002282	мж 0,000084	фа 0,000245	ьл 0,000852	
кб 0,000005	на 0,009777	тв 0,001752	ьн 0,001384	ек 0,002468	мз 0,000189	уя 0,000286	ьм 0,000080	
кв 0,000198	кв 0,000453	тг 0,000029	ьо 0,000034	ел 0,006831	ми 0,004256	фд 0,000004	ьн 0,000510	
кг 0,000003	нб 0,000017	тд 0,000122	ьп 0,000029	ем 0,005149	мк 0,000562	фе 0,000307	ьп 0,000072	
кд 0,000024	нв 0,000062	те 0,005023	ьр 0,000005	ен 0,010611	мл 0,000274	фи 0,000321	ьр 0,000093	
ке 0,001134	нг 0,000594	тж 0,000002	ьт 0,001017	ео 0,001476	мм 0,000508	фк 0,000014	ьс 0,000031	
кж 0,000047	нд 0,000422	тн 0,003617	ьу 0,000153	еп 0,003217	мн 0,003083	фл 0,000074	ьт 0,002723	
кз 0,000016	не 0,009083	тк 0,000836	ьу 0,000005	ер 0,008576	мо 0,005715	фм 0,000002	ьф 0,000045	
ки 0,002638	нк 0,000034	тл 0,000212	ьф 0,000017	ес 0,007121	мп 0,001208	фн 0,000010	ьх 0,000004	
кк 0,000050	нз 0,000053	тм 0,000090	ьч 0,000138	ет 0,006909	мр 0,000294	фо 0,000465	ьш 0,000006	
кл 0,000896	ни 0,006827	тн 0,001134	ьц 0,000091	еу 0,000729	мс 0,000992	фп 0,000010	ьз 0,000004	
км 0,000047	нк 0,000589	то 0,013735	ьщ 0,000660	еф 0,000181	мт 0,000395	фр 0,000463	юа 0,000041	
кн 0,000414	нл 0,000076	тп 0,000202	ьз 0,000005	ех 0,001002	мү 0,002406	фс 0,000002	ья 0,000006	
ко 0,008102	нм 0,000002	тр 0,000260	ьз 0,000005	ец 0,000354	мф 0,000056	фт 0,000037	юб 0,000372	
кп 0,000002	нн 0,002595	тс 0,001151	ьо 0,000615	еч 0,001799	мх 0,000074	фү 0,000167	юв 0,000274	
кр 0,001780	но 0,009443	тт 0,000283	ья 0,000281	еш 0,001229	мц 0,000037	фф 0,000066	юг 0,000099	
кс 0,000289	нп 0,000005	тү 0,001465	ьб 0,000005	ещ 0,000772	мч 0,000327	фч 0,000004	юд 0,000626	
кт 0,000638	нр 0,000134	тф 0,000043	ьв 0,000010	ез 0,000288	мш 0,000058	фы 0,000021	юе 0,000072	
ку 0,002173	нс 0,000601	тх 0,000009	ьг 0,000007	ею 0,000136	миц 0,000016	фь 0,000010	юж 0,000115	
кф 0,000002	нт 0,000972	тц 0,000090	эд 0,000062	жа 0,001521	мы 0,000897	ха 0,000743	юз 0,000068	
кх 0,000002	нү 0,002795	тч 0,000124	эе 0,000002	яа 0,000784	мь 0,000253	фя 0,000002	юи 0,000230	
кц 0,000033	нф 0,000081	тш 0,000009	ээ 0,000003	жб 0,000025	мэ 0,000274	хб 0,000115	юй 0,000037	
кч 0,000003	нх 0,000047	тщ 0,000007	эй 0,000059	жв 0,000033	мю 0,000010	хв 0,000584	юк 0,000331	
кш 0,000002	нц 0,000286	ть 0,000010	эк 0,000079	жг 0,000010	на 0,011708	хг 0,000088	юл 0,000093	
кщ 0,000009	нч 0,000240	ты 0,001728	эл 0,000713	жд 0,000774	мя 0,000829	хд 0,000167	юм 0,000216	
кь 0,000009	нш 0,000016	тэ 0,005576	эм 0,000067	же 0,003505	нб 0,000144	хе 0,000105	юн 0,000370	
кэ 0,000017	нщ 0,000128	тз 0,000005	эн 0,000427	жж 0,000004	нв 0,000438	хж 0,000033	юо 0,000208	
кю 0,000009	ны 0,002931	тю 0,000047	эп 0,000053	жз 0,000012	нг 0,000766	хз 0,000093	юп 0,000327	
ла 0,005461	нь 0,001249	тя 0,000644	эр 0,000078	жи 0,001817	нд 0,000650	хи 0,000385	юр 0,000196	
ля 0,000007	нэ 0,000005	уа 0,000155	эс 0,000024	жк 0,000228	не 0,010903	хк 0,000224	юс 0,000568	
лб 0,000047	ню 0,000210	уб 0,000782	эт 0,002278	жл 0,000023	нж 0,000070	хл 0,000270	ют 0,000545	
лв 0,000022	оа 0,001894	ув 0,000775	эф 0,000038	жм 0,000025	нз 0,000206	хм 0,000204	юу 0,000058	
лг 0,000133	об 0,000017	ур 0,000073	эк 0,000003	жн 0,001039	ни 0,008403	хн 0,000506	юф 0,000031	
лд 0,000078	об 0,003476	уд 0,001525	эш 0,000005	жо 0,000366	нк 0,000840	хо 0,002651	юх 0,000064	
ле 0,004399	ов 0,006114	уе 0,000134	ээ 0,000003	жп 0,000037	нл 0,000134	хп 0,000327	юц 0,000025	
лж 0,000219	ог 0,004761	уж 0,001234	юа 0,000005	жр 0,000008	нм 0,000158	хр 0,000286	юш 0,000428	
лз 0,000048	од 0,005101	уз 0,000240	яб 0,000005	жс 0,000039	нн 0,003488	хс 0,000319	ют 0,000066	
ли 0,006679	ое 0,001806	уи 0,000176	юб 0,000246	жт 0,000025	но 0,011512	ху 0,000150	юч 0,000613	
лк 0,000405	ок 0,002026	уй 0,000105	юв 0,000012	жү 0,000278	нп 0,000576	хү 0,000243	юз 0,000062	
лл 0,000789	оз 0,001270	ул 0,000813	юг 0,000028	жч 0,000158	нр 0,000245	хф 0,000023	юя 0,000049	
лм 0,000053	оа 0,000619	ум 0,000200	юд 0,000365	жь 0,000037	нс 0,001033	хх 0,000041	юи 0,000115	
лн 0,000352	ой 0,004132	ун 0,001232	юе 0,000003	жз 0,000012	нт 0,001290	хц 0,000012	яа 0,000183	
ло 0,005499	ок 0,002168	ун 0,000345	юж 0,000031	за 0,005657	нү 0,003497	хч 0,000086	яб 0,000622	

Оцінка ентропії

[illegible]

Надлишковість для H1 3.189340 --> 0.9377738

Надлишковість для H2 3.420796 --> 0.9321862

[illegible]

Надлишковість для H1 1.754870 --> 0.9826057

Надлишковість для H2 2.421119--> 0.9760018

[illegible]

Надлишковість для H1 1.045372 --> 0.9930922

Надлишковість для H2 1.573794--> 0.9896003

Код

```
package com.company;

import java.io.*;
import java.util.HashMap;
import java.util.Map;
import java.util.Scanner;

public class Main {

    public static void main(String[] args) throws IOException {

        String path = ("C:\\Users\\kateh\\Documents\\кпи\\крипта\\king1.txt");
        String str = ReadText(path);

        String Capital = CapitalLetters(str);

        String Remove = RemoveSpace(Capital);
        String Leave = LeaveSpace(Capital);

        int AmountR = AmountLetters(Remove);
        int AmountL = AmountLetters(Leave);

        //      System.out.println("Monogramma");
        //
        //      System.out.println("Без пробелов");
        //      //      Monogram(Remove, AmountR);
        //      float count1 = (float) ((1 - (Monogram(Remove, AmountR)) / (Math.Log(33) / Math.Log(2))));
        //      System.out.println("Redundancy = " + count1 + "\n");
        //
        //
        //      System.out.println("С пробелами");
        //      //      Monogram(Leave, AmountL);
        //      float count2 = (float) ((1 - (Monogram(Leave, AmountL)) / (Math.Log(33) / Math.Log(2))));
        //      System.out.println("Redundancy = " + count2 + "\n");

        System.out.println("BigrammaPut");

        System.out.println("с пробелами");
        //BigramWr(Leave, AmountL);
        float a = (float) ((1 - (BigramWr(Leave, AmountL)) / (Math.Log(33) / Math.Log(2))));
        System.out.println("Redundancy = " + a + "\n");

        System.out.println("без");
        //BigramWr(Remove, AmountR);
        float b = (float) ((1 - (BigramWr(Remove, AmountR)) / (Math.Log(33) / Math.Log(2))));
        System.out.println("Redundancy = " + b + "\n");

        System.out.println("Bigramma");

        System.out.println("с пробелами");
        BigramWithOut(Leave, AmountL);
        float c = (float) ((1 - (BigramWr(Leave, AmountL)) / (Math.Log(33) / Math.Log(2))));
        System.out.println("Redundancy = " + c + "\n");

        System.out.println("без");
        BigramWithOut(Remove, AmountR);
        float d = (float) ((1 - (BigramWr(Remove, AmountR)) / (Math.Log(33) / Math.Log(2))));
        System.out.println("Redundancy = " + d + "\n");

    }

    public static String ReadText(String path) {
        String text = "";
        String buf = "";
        try {
            FileInputStream File = new FileInputStream(path);
            DataInputStream dis = new DataInputStream(File);
            BufferedReader br = new BufferedReader(new InputStreamReader(dis));
        }
    }
}
```

```

        while ((buf = br.readLine()) != null) {
            text += buf;
        }
    } catch (Exception e) {
        e.printStackTrace();
    }
    return text;
}

public static String CapitalLetters(String text) {
    String str = text.toLowerCase();
    return str;
}

public static String RemoveSpace(String text) {
    String str1 = text.replaceAll("[^a-я]", "");
    return str1;
}

public static String LeaveSpace(String text) {
    String str1 = text.replaceAll("[^a-я ]", "");
    return str1;
}

public static int AmountLetters(String text) {
    int Amount = text.length();
    return Amount;
}

public static float Monogram(String text, int amount) {
    char[] strchar = text.toCharArray();
    Map<Character, Float> MonoMap = new HashMap<Character, Float>();
    int calc = 0;
    char letter = strchar[calc];
    for (int i = 0; i < text.length(); i++) {
        letter = strchar[i];
        if (!MonoMap.containsKey(letter)) {
            MonoMap.put(letter, (float) 1);
        } else {
            MonoMap.put(letter, MonoMap.compute(letter, (k, v) -> (float) v + 1));
        }
    }
    float res = 0;
    for (Character key : MonoMap.keySet()) {
        float a = MonoMap.get(key) / amount;
        System.out.format(key + " " + "%.6f%n", a);
    }
    for (float values : MonoMap.values()) {
        res += (float) ((values / amount) * (Math.Log(values / amount)) / Math.Log(2));
    }
    System.out.println("Result = " + (-res) + "\n");
    // return 0;
    return (-res);
}

public static float BigramWr(String str, int amount) {
    char[] strtchar = str.toCharArray();
    Character ch1, ch2;
    Map<String, Float> BiMap = new HashMap<String, Float>();
    for (int i = 0; i < str.length() - 1; i++) {
        ch1 = strtchar[i];
        ch2 = strtchar[i + 1];
        String s1 = ch1.toString();
        String s2 = ch2.toString();
        String s3 = s1 + s2;
        if (!BiMap.containsKey(s3)) {
            BiMap.put(s3, (float) 1);
        } else {
            BiMap.put(s3, BiMap.compute(s3, (k, v) -> (float) v + 1));
        }
    }
    float res = 0;
    for (String key : BiMap.keySet()) {
        System.out.format(key + " " + "%.6f%n", BiMap.get(key)/amount);
    }
    for (float values : BiMap.values()) {
        res += (float) ((values / amount) * (Math.Log(values / amount)) / Math.Log(2));
    }
}

```

```

        System.out.println("Result = " + (-res/2) + " \n");
        // return 0;
        return (-res / 2);
    }

    public static float BigramWithOut(String str, int amount) {
        Map<String,Float> BiMap = new HashMap<String, Float>();
        char [] chtostr = str.toCharArray();
        Character ch1,ch2;
        for (int i=0; i<str.length()-1;i++){
            ch1= chtostr[i];
            ch2 = chtostr[i+1];
            String s1 = ch1.toString();
            String s2 = ch2.toString();
            String s3 = s1 + s2;
            if(BiMap.containsKey(s3)){
                BiMap.put(s3, BiMap.compute(s3, (k,v)-> (float) v+1));
            }
            else {
                BiMap.put(s3,(float) 1);
            }

            i++;
        }
        for (String key : BiMap.keySet() ){
            System.out.format(key + " " + "%.6f\n",BiMap.get(key)/amount);
        }
        float res = 0;
        for (float values : BiMap.values()){
            res += ((values/amount)*((Math.Log(values/amount))/Math.Log(2)));
        }
        System.out.println("Result = " + (-res/2));
        return (-res/2);
    }
}

```

Висновок

В цій лабораторній роботі ми навчилися оцінювати ентропію та знаходити надлишковість. Порівнювали різні моделі джерел відкритого тексту для наближеного визначення ентропії.