



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-71
Рейценштейн Кирило і Таран Вікторія
Перевірили:
Чорний О.
Савчук М. М.
Завадська Л. О.

Київ 2019

Мета роботи :

Засвоєння понять ентропії на символ джерела та його надлишковості, вивчення та порівняння різних моделей джерела відкритого тексту для наближеного визначення ентропії, набуття практичних навичок щодо оцінки ентропії на символ джерела.

Порядок виконання роботи:

0. Уважно прочитати методичні вказівки до виконання комп'ютерного практикуму.
1. Написати програми для підрахунку частот букв і частот біграм в тексті, а також підрахунку H_1 та H_2 за безпосереднім означенням. Підрахувати частоти букв та біграм, а також значення H_1 та H_2 на довільно обраному тексті російською мовою достатньої довжини (щонайменше 1Мб), де імовірності замінити відповідними частотами. Також одержати значення H_1 та H_2 на тому ж тексті, в якому вилучено всі пробіли.
2. За допомогою програми CoolPinkProgram оцінити значення H^{10} , H^{20} , H^{30} .
3. Використовуючи отримані значення ентропії, оцінити надлишковість російської мови в різних моделях джерела.

Результати:

Таблиця частот монограм, ентропія та надлишковість.

З пробілом	Без пробіла
а: 0.0693081 б: 0.0146276 в: 0.0393918 г: 0.0152046 д: 0.0246376 е: 0.0731772 ж: 0.0094803 з: 0.0136825 и: 0.0555528 й: 0.0087950 к: 0.0286777 л: 0.0419655 м: 0.0239848 н: 0.0580773 о: 0.0961128 п: 0.0201761 р: 0.0333131 с: 0.0444628 т: 0.0500960 у: 0.0225640 ф: 0.0010540 х: 0.0065013 ц: 0.0023630 ч: 0.0141234 ш: 0.0071422 щ: 0.0023991 ы: 0.0155129 ь: 0.0167326 э: 0.0029701 ю: 0.0052148 я: 0.0180209 : 0.1646782	а: 0.0829717 б: 0.0175114 в: 0.0471577 г: 0.0182021 д: 0.0294947 е: 0.0876035 ж: 0.0113493 з: 0.0163800 и: 0.0665046 й: 0.0105289 к: 0.0343313 л: 0.0502387 м: 0.0287133 н: 0.0695269 о: 0.1150608 п: 0.0241537 р: 0.0398805 с: 0.0532284 т: 0.0599720 у: 0.0270123 ф: 0.0012617 х: 0.0077830 ц: 0.0028288 ч: 0.0169078 ш: 0.0085502 щ: 0.0028720 ы: 0.0185712 ь: 0.0200313 э: 0.0035557 ю: 0.0062428 я: 0.0215736
$H_1 = 4.3536432$	$H_{1_ns} = 4.4393172$
$R_1 = 0.1292714$	$R_{1_ns} = 0.1039279$

Таблиця частот біграм з пробілом, що перетинаються, ентропія та надлишковість.

Від А до О, від Р до Я та пробіл.

	а	б	в	г	д	е	ж	з	и	й	к	л	м	н	о	п
а	0.0000221	0.0005655	0.0029298	0.0005615	0.0020737	0.0009791	0.0010680	0.0039701	0.0001303	0.0005014	0.0039871	0.0091538	0.0022924	0.0044285	0.0000391	0.0005728

б	0.0007932	0.0000011	0.0000816	0.0000051	0.0000306	0.0018947	0.0000051	0.0000023	0.0007371	0.0000000	0.0002266	0.0007621	0.0000544	0.0002850	0.0018737	0.0000000
в	0.0052013	0.0000159	0.0000238	0.0000300	0.0002963	0.0040647	0.0000011	0.0005060	0.0043004	0.0000000	0.0001456	0.0007405	0.0001371	0.0012051	0.0066710	0.0001960
г	0.0007955	0.0000006	0.0000074	0.0000028	0.0010884	0.0003394	0.0000045	0.0000006	0.0007133	0.0000000	0.0001128	0.0014255	0.0000006	0.0001643	0.0079657	0.0000011
д	0.0040132	0.0000210	0.0007230	0.0000040	0.0000368	0.0042772	0.0000130	0.0000011	0.0019417	0.0000000	0.0001977	0.0007553	0.0000657	0.0014465	0.0035083	0.0000629
е	0.0000606	0.0012255	0.0024018	0.0034749	0.0023281	0.0020516	0.0007496	0.0011077	0.0001054	0.0023763	0.0016794	0.0054738	0.0038069	0.0069651	0.0002130	0.0012947
ж	0.0013156	0.0000414	0.0000000	0.0000051	0.0007303	0.0038545	0.0000108	0.0000000	0.0013876	0.0000000	0.0000850	0.0000034	0.0000023	0.0009405	0.0000419	0.0000000
з	0.0050653	0.0001513	0.0007643	0.0004901	0.0007281	0.0002249	0.0000992	0.0000555	0.0002867	0.0000000	0.0000895	0.0001745	0.0002629	0.0018460	0.0003405	0.0000204
и	0.0000975	0.0003666	0.0025485	0.0003972	0.0014748	0.0022947	0.0002969	0.0017020	0.0005893	0.0012680	0.0015128	0.0042092	0.0026477	0.0039293	0.0001088	0.0001462
й	0.0000057	0.0000068	0.0000164	0.0000000	0.0001977	0.0000028	0.0000000	0.0000011	0.0000062	0.0000000	0.0000674	0.0000895	0.0000533	0.0002714	0.0000125	0.0000091
к	0.0073402	0.0000000	0.0001632	0.0000000	0.0000057	0.0004759	0.0000136	0.0000040	0.0027179	0.0000000	0.0000181	0.0003128	0.0000011	0.0005910	0.0074467	0.0000023
л	0.0072405	0.0000164	0.0000108	0.0001105	0.0000142	0.0049809	0.0003570	0.0000034	0.0050155	0.0000000	0.0002731	0.0002436	0.0000040	0.0002861	0.0055073	0.0000442
м	0.0023576	0.0000136	0.0000193	0.0000289	0.0000045	0.0028165	0.0000011	0.0000040	0.0025083	0.0000000	0.0000555	0.0000924	0.0000431	0.0012136	0.0029967	0.0000601
н	0.0104394	0.0000130	0.0000147	0.0000980	0.0006040	0.0093958	0.0000040	0.0000136	0.0073691	0.0000000	0.0002448	0.0000045	0.0000028	0.0027723	0.0092547	0.0000051
о	0.0000181	0.0032721	0.0069957	0.0043882	0.0040007	0.0023610	0.0018006	0.0009694	0.0006918	0.0029134	0.0014907	0.0052291	0.0045979	0.0079487	0.0002045	0.0010913
п	0.0010918	0.0000068	0.0000000	0.0000000	0.0000000	0.0019916	0.0000000	0.0000000	0.0007706	0.0000000	0.0000555	0.0004346	0.0000000	0.0000612	0.0077629	0.0000227
р	0.0065390	0.0002147	0.0002895	0.0003400	0.0003207	0.0049990	0.0002844	0.0000147	0.0041877	0.0000000	0.0005360	0.0000487	0.0001258	0.0006323	0.0065651	0.0000567
с	0.0016930	0.0000493	0.0017281	0.0000102	0.0002992	0.0041684	0.0000244	0.0000176	0.0012573	0.0000000	0.0048036	0.0024046	0.0007762	0.0008997	0.0022913	0.0015916
т	0.0047191	0.0000147	0.0025372	0.0000062	0.0001145	0.0047554	0.0000045	0.0000017	0.0036341	0.0000000	0.0004584	0.0002527	0.0000215	0.0010250	0.0136905	0.0000499
у	0.0000436	0.0004969	0.0010692	0.0009224	0.0014454	0.0002436	0.0016204	0.0002816	0.0000125	0.0001128	0.0007241	0.0010754	0.0011083	0.0002153	0.0000255	0.0005496
ф	0.0000844	0.0000000	0.0000000	0.0000000	0.0000000	0.0001162	0.0000000	0.0000000	0.0002952	0.0000000	0.0000000	0.0000232	0.0000062	0.0000017	0.0000929	0.0000000
х	0.0007768	0.0000017	0.0000935	0.0000017	0.0000034	0.0000334	0.0000000	0.0000000	0.0001082	0.0000000	0.0000040	0.0000833	0.0000465	0.0000963	0.0021055	0.0000068
ц	0.0004448	0.0000028	0.0000448	0.0000023	0.0000028	0.0006374	0.0000006	0.0000017	0.0001847	0.0000000	0.0001099	0.0000006	0.0000051	0.0000062	0.0002691	0.0000096
ч	0.0021196	0.0000000	0.0000057	0.0000006	0.0000000	0.0032007	0.0000000	0.0000006	0.0012278	0.0000000	0.0002958	0.0000419	0.0000045	0.0004646	0.0000465	0.0000000
ш	0.0008561	0.0000000	0.0000793	0.0000000	0.0000000	0.0021400	0.0000000	0.0000000	0.0014386	0.0000000	0.0003360	0.0005116	0.0000045	0.0003473	0.0002635	0.0000057
щ	0.0002748	0.0000000	0.0000000	0.0000000	0.0000000	0.0012414	0.0000000	0.0000000	0.0006312	0.0000000	0.0000000	0.0000000	0.0000000	0.0000300	0.0000011	0.0000000
ы	0.0000017	0.0004402	0.0009349	0.0000850	0.0001309	0.0009020	0.0000238	0.0000589	0.0000102	0.0011303	0.0001575	0.0024414	0.0010998	0.0003173	0.0000045	0.0001105
ь	0.0000130	0.0000884	0.0000414	0.0000510	0.0000317	0.0005762	0.0000011	0.0001603	0.0003955	0.0000000	0.0011343	0.0000136	0.0002091	0.0011320	0.0000312	0.0000130
э	0.0000000	0.0000011	0.0000011	0.0000023	0.0000017	0.0000000	0.0000000	0.0000057	0.0000000	0.0000028	0.0000374	0.0000402	0.0000017	0.0000227	0.0000000	0.0000204
ю	0.0000074	0.0005961	0.0000085	0.0000051	0.0002283	0.0000006	0.0000068	0.0000040	0.0000023	0.0000000	0.0000091	0.0000176	0.0000153	0.0000312	0.0000068	0.0000108
я	0.0000040	0.0000249	0.0002646	0.0002635	0.0005111	0.0000731	0.0002720	0.0002986	0.0000436	0.0001133	0.0001094	0.0007757	0.0003071	0.0006408	0.0000159	0.0000584
	0.0029196	0.0063566	0.0139172	0.0032709	0.0068852	0.0049690	0.0024142	0.0039140	0.0104882	0.0000023	0.0085000	0.0033440	0.0052602	0.0153875	0.0126650	0.0133058
р	с	т	у	ф	х	ц	ч	ш	щ	ы	ь	э	ю	я		
а	0.0022352	0.0033633	0.0039758	0.0000521	0.0002442	0.0008034	0.0001145	0.0007473	0.0006046	0.0001926	0.0000000	0.0000000	0.0000079	0.0007961	0.0022902	0.0176527
б	0.0010414	0.0000759	0.0000074	0.0009241	0.0000000	0.0000516	0.0000057	0.0000057	0.0000079	0.0002357	0.0039338	0.0001201	0.0000000	0.0000045	0.0004737	0.0003649
в	0.0010102	0.0032925	0.0002187	0.0006255	0.0000000	0.0000544	0.0000221	0.0000748	0.0010097	0.0000074	0.0022732	0.0001824	0.0000000	0.0000023	0.0002482	0.0051588
г	0.0006476	0.0000289	0.0000062	0.0005575	0.0000011	0.0000011	0.0000000	0.0000357	0.0000051	0.0000000	0.0000000	0.0000000	0.0000000	0.0000034	0.0000000	0.0006482
д	0.0014040	0.0003218	0.0001768	0.0017003	0.0000000	0.0000578	0.0001660	0.0000255	0.0000963	0.0000006	0.0005026	0.0007853	0.0000006	0.0000068	0.0003734	0.0009037
е	0.0049679	0.0036755	0.0043390	0.0000907	0.0000102	0.0007451	0.0001524	0.0008578	0.0008805	0.0005774	0.0000000	0.0000000	0.0000011	0.0003745	0.0003349	0.0177405
ж	0.0000006	0.0000901	0.0000000	0.0002295	0.0000000	0.0000000	0.0000006	0.0000380	0.0000000	0.0000000	0.0000000	0.0000385	0.0000000	0.0000011	0.0000000	0.0002601
з	0.0001926	0.0000499	0.0000125	0.0002465	0.0000000	0.0000000	0.0000040	0.0000521	0.0000062	0.0000000	0.0003609	0.0001643	0.0000000	0.0000040	0.0004499	0.0009581
и	0.0005116	0.0020482	0.0036318	0.0000119	0.0000130	0.0012420	0.0007740	0.0017865	0.0003717	0.0001405	0.0000000	0.0000000	0.0000034	0.0002363	0.0016482	0.0171796
й	0.0000045	0.0004788	0.0002431	0.0000006	0.0000040	0.0000000	0.0000397	0.0000686	0.0000493	0.0000011	0.0000000	0.0000000	0.0000011	0.0000000	0.0000006	0.0067895
к	0.0013122	0.0008816	0.0003932	0.0013372	0.0000011	0.0000011	0.0000068	0.0000000	0.0000102	0.0000000	0.0000000	0.0000000	0.0000000	0.0000000	0.0000000	0.0044211
л	0.0000017	0.0011507	0.0000249	0.0011275	0.0000040	0.0000000	0.0000006	0.0001105	0.0000034	0.0000006	0.0008822	0.0033503	0.0000006	0.0009547	0.0016465	0.0068138
м	0.0000708	0.0001156	0.0000538	0.0025961	0.0000011	0.0000000	0.0000091	0.0000204	0.0000040	0.0000045	0.0008788	0.0000674	0.0000040	0.0000028	0.0003672	0.0065532
н	0.0000822	0.0010097	0.0004011	0.0023060	0.0000153	0.0000011	0.0003292	0.0002425	0.0000011	0.0001501	0.0027621	0.0011672	0.0000017	0.0001462	0.0016142	0.0051390
о	0.0051764	0.0056036	0.0062087	0.0000561	0.0001552	0.0003247	0.0000850	0.0015207	0.0011723	0.0001399	0.0000000	0.0000000	0.0000453	0.0007479	0.0004742	0.0223384
п	0.0058925	0.0000079	0.0000459	0.0004861	0.0000011	0.0000000	0.0000136	0.0000295	0.0000125	0.0000028	0.0002380	0.0000334	0.0000006	0.0000028	0.0003173	0.0000357
р	0.0000164	0.0001972	0.0004357	0.0021457	0.0000102	0.0000907	0.0000266	0.0000640	0.0001989	0.0000289	0.0011615	0.0008142	0.0000000	0.0001060	0.0008357	0.0006091
с	0.0001150	0.0005779	0.0100071	0.0004890	0.0000074	0.0001275	0.0000550	0.0004074	0.0000578	0.0000017	0.0003116	0.0026800	0.0000017	0.0001179	0.0028471	0.0027514
т	0.0023060	0.0008142	0.0000793	0.0011009	0.0000204	0.0000096	0.0000572	0.0002601	0.0000006	0.0000119	0.0014006	0.0059350	0.0000034	0.0000442	0.0004963	0.0041384

у	0.0005252	0.0009609	0.0010006	0.0000057	0.0000334	0.0002788	0.0000034	0.0008233	0.0006578	0.0001864	0.0000000	0.0000000	0.0000147	0.0009088	0.0001094	0.0061486
ф	0.0002374	0.0000102	0.0000176	0.0000238	0.0000068	0.0000000	0.0000000	0.0000006	0.0000000	0.0000000	0.0000045	0.0000504	0.0000000	0.0000011	0.0000011	0.0000357
х	0.0000640	0.0000402	0.0000108	0.0001422	0.0000000	0.0000000	0.0000000	0.0000006	0.0000085	0.0000000	0.0000000	0.0000006	0.0000006	0.0000000	0.0000006	0.0025956
ц	0.0000023	0.0000062	0.0000051	0.0002011	0.0000000	0.0000000	0.0000074	0.0000017	0.0000011	0.0000000	0.0001422	0.0000000	0.0000057	0.0000000	0.0000017	0.0001654
ч	0.0000567	0.0000057	0.0041678	0.0008607	0.0000000	0.0000000	0.0000000	0.0000000	0.0001581	0.0000000	0.0000000	0.0001513	0.0000000	0.0000000	0.0000000	0.0007139
ш	0.0000011	0.0000000	0.0000601	0.0002680	0.0000011	0.0000000	0.0000023	0.0000000	0.0000017	0.0000000	0.0000000	0.0004572	0.0000000	0.0000034	0.0000000	0.0000601
щ	0.0000051	0.0000000	0.0000000	0.0000839	0.0000000	0.0000000	0.0000000	0.0000000	0.0000000	0.0000000	0.0000000	0.0000221	0.0000000	0.0000000	0.0000000	0.0000074
ы	0.0002623	0.0007298	0.0006754	0.0000062	0.0000006	0.0007859	0.0000006	0.0001677	0.0003989	0.0000006	0.0000000	0.0000000	0.0000000	0.0000000	0.0000023	0.0039735
ь	0.0000034	0.0007910	0.0000720	0.0000028	0.0000028	0.0000011	0.0000969	0.0000589	0.0003847	0.0000142	0.0000000	0.0000000	0.0000045	0.0003728	0.0004992	0.0098241
э	0.0000068	0.0000040	0.0026907	0.0000000	0.0000017	0.0000000	0.0000000	0.0000000	0.0000006	0.0000000	0.0000000	0.0000000	0.0000000	0.0000000	0.0000000	0.0000028
ю	0.0000380	0.0001643	0.0002470	0.0000017	0.0000017	0.0000102	0.0000102	0.0000408	0.0000238	0.0003139	0.0000000	0.0000000	0.0000011	0.0000323	0.0000000	0.0031582
я	0.0000453	0.0008091	0.0012986	0.0000051	0.0000000	0.0001790	0.0000329	0.0001218	0.0000453	0.0001904	0.0000000	0.0000000	0.0000017	0.0000941	0.0001275	0.0105273
	0.0036585	0.0152657	0.0074569	0.0039191	0.0004725	0.0014595	0.0002470	0.0059600	0.0006652	0.0000958	0.0000006	0.0000006	0.0027440	0.0000289	0.0020947	0.0000000
H2 = 3.8126644																
R2 = 0.2374671																

Таблиця частот біграм без пробіла, що перетинаються, ентропія та надлишковість.
Від А до П, від Р до Я.

	а	б	в	г	д	е	ж	з	и	й	к	л	м	н	о	п
а	0.0004992	0.0014452	0.0053932	0.0010980	0.0033795	0.0020278	0.0014506	0.0051503	0.0015448	0.0005954	0.0060492	0.0113186	0.0033311	0.0073840	0.0019276	0.0023515
б	0.0009581	0.0000034	0.0001076	0.0000087	0.0000383	0.0022674	0.0000061	0.0000047	0.0009170	0.0000000	0.0002732	0.0009076	0.0000680	0.0003492	0.0023494	0.0000067
в	0.0062531	0.0001743	0.0003216	0.0003176	0.0007253	0.0050777	0.0000612	0.0006964	0.0053333	0.0000000	0.0006816	0.0009978	0.0003647	0.0018563	0.0083536	0.0008195
г	0.0009554	0.0000397	0.0000949	0.0000135	0.0013503	0.0004292	0.0000141	0.0000249	0.0008928	0.0000000	0.0001588	0.0017022	0.0000188	0.0002698	0.0095074	0.0001097
д	0.0047870	0.0000706	0.0009332	0.0000283	0.0000774	0.0051106	0.0000289	0.0000383	0.0023871	0.0000000	0.0003081	0.0009177	0.0001056	0.0018549	0.0042272	0.0001433
е	0.0003277	0.0024699	0.0047574	0.0045549	0.0037852	0.0027787	0.0011613	0.0020844	0.0015259	0.0028217	0.0028695	0.0070086	0.0057525	0.0099669	0.0017769	0.0034878
ж	0.0015656	0.0000619	0.0000209	0.0000074	0.0008827	0.0045912	0.0000141	0.0000094	0.0016679	0.0000000	0.0001144	0.0000081	0.0000108	0.0011566	0.0000727	0.0000316
з	0.0060263	0.0002039	0.0009978	0.0006163	0.0009345	0.0002987	0.0001305	0.0000881	0.0003822	0.0000000	0.0002079	0.0002308	0.0003835	0.0022721	0.0004582	0.0001325
и	0.0003445	0.0012851	0.0050931	0.0008955	0.0026367	0.0033055	0.0005699	0.0025809	0.0017607	0.0015071	0.0028762	0.0054046	0.0037590	0.0066036	0.0018717	0.0022034
й	0.0004017	0.0002866	0.0006789	0.0002092	0.0006863	0.0001191	0.0001769	0.0001702	0.0007542	0.0000000	0.0005598	0.0002738	0.0003377	0.0009117	0.0005362	0.0007138
к	0.0087989	0.0003734	0.0005450	0.0000895	0.0002227	0.0006782	0.0002348	0.0000962	0.0035235	0.0000000	0.0003075	0.0004757	0.0001931	0.0013611	0.0092679	0.0003263
л	0.0087270	0.0002072	0.0008053	0.0002792	0.0002436	0.0064340	0.0004629	0.0001366	0.0064239	0.0000000	0.0007852	0.0006049	0.0001561	0.0009379	0.0076034	0.0005665
м	0.0029280	0.0002900	0.0006755	0.0002503	0.0003586	0.0034663	0.0001359	0.0001837	0.0036009	0.0000000	0.0005295	0.0002779	0.0002590	0.0020958	0.0041902	0.0007488
н	0.0127322	0.0003371	0.0006123	0.0002267	0.0009426	0.0112540	0.0000653	0.0002220	0.0090983	0.0000000	0.0004885	0.0000928	0.0001352	0.0040180	0.0113253	0.0006997
о	0.0002839	0.0052398	0.0107528	0.0056879	0.0058151	0.0036405	0.0026973	0.0017594	0.0022189	0.0034596	0.0028776	0.0067207	0.0062652	0.0120580	0.0027726	0.0033903
п	0.0012978	0.0000101	0.0000040	0.0000000	0.0000013	0.0023663	0.0000020	0.0000013	0.0009211	0.0000000	0.0000693	0.0005167	0.0000013	0.0000747	0.0092215	0.0000276
р	0.0077810	0.0002833	0.0003976	0.0004212	0.0004084	0.0059503	0.0003411	0.0000357	0.0050582	0.0000000	0.0006977	0.0000659	0.0001669	0.0008222	0.0078590	0.0001292
с	0.0020776	0.0001615	0.0022801	0.0001009	0.0005234	0.0050178	0.0002160	0.0000834	0.0015925	0.0000000	0.0059563	0.0029274	0.0010422	0.0015010	0.0029025	0.0021092
т	0.0056960	0.0002833	0.0034602	0.0000774	0.0003451	0.0058480	0.0000949	0.0001023	0.0046450	0.0000000	0.0007845	0.0004003	0.0002039	0.0017083	0.0165577	0.0004212
у	0.0002106	0.0008006	0.0018617	0.0012137	0.0020406	0.0004346	0.0020258	0.0004837	0.0007118	0.0001346	0.0013402	0.0014075	0.0015643	0.0011613	0.0005073	0.0012070
ф	0.0001070	0.0000034	0.0000094	0.0000007	0.0000000	0.0001379	0.0000000	0.0000000	0.0003539	0.0000000	0.0000020	0.0000276	0.0000087	0.0000061	0.0001110	0.0000020
х	0.0009715	0.0001225	0.0003532	0.0000881	0.0001823	0.0001023	0.0000518	0.0000693	0.0003922	0.0000000	0.0001763	0.0002166	0.0001675	0.0003512	0.0027572	0.0003041
ц	0.0005342	0.0000121	0.0000733	0.0000081	0.0000114	0.0007623	0.0000027	0.0000074	0.0002362	0.0000000	0.0001426	0.0000013	0.0000101	0.0000269	0.0003351	0.0000309
ч	0.0025284	0.0000256	0.0001198	0.0000128	0.0000343	0.0038108	0.0000054	0.0000350	0.0015084	0.0000000	0.0003828	0.0000599	0.0000168	0.0006338	0.0001164	0.0001076
ш	0.0010180	0.0000061	0.0000989	0.0000020	0.0000007	0.0025412	0.0000007	0.0000013	0.0017136	0.0000000	0.0004030	0.0006089	0.0000114	0.0004158	0.0003162	0.0000155
щ	0.0003263	0.0000000	0.0000013	0.0000000	0.0000000	0.0014741	0.0000000	0.0000000	0.0007495	0.0000000	0.0000000	0.0000013	0.0000007	0.0000363	0.0000020	0.0000020
ы	0.0000720	0.0007569	0.0015286	0.0001992	0.0003768	0.0011868	0.0000861	0.0002436	0.0004097	0.0013422	0.0004205	0.0029785	0.0014660	0.0009655	0.0002927	0.0005739
ь	0.0001601	0.0004023	0.0011384	0.0002200	0.0004777	0.0013315	0.0001110	0.0004844	0.0013550	0.0000000	0.0020251	0.0002234	0.0005981	0.0024241	0.0009285	0.0009238
э	0.0000000	0.0000027	0.0000020	0.0000027	0.0000020	0.0000000	0.0000000	0.0000067	0.0000000	0.0000034	0.0000444	0.0000478	0.0000020	0.0000283	0.0000000	0.0000242
ю	0.0000976	0.0007946	0.0002577	0.0000908	0.0004266	0.0001050	0.0001265	0.0000908	0.0003324	0.0000000	0.0002422	0.0001130	0.0001426	0.0003055	0.0003129	0.0002920
я	0.0003256	0.0004777	0.0014075	0.0005658	0.0011014	0.0005187	0.0005046	0.0006654	0.0010502	0.0001352	0.0008289	0.0011734	0.0007260	0.0024719	0.0008121	0.0010368
	р	с	т	у	ф	х	ц	ч	ш	щ	ы	ь	э	ю	я	
а	0.0030720	0.0059651	0.0056684	0.0005652	0.0003411	0.0011041	0.0001601	0.0015515	0.0008000	0.0002368	0.0000000	0.0000000	0.0003922	0.0009513	0.0030391	
б	0.0012373	0.0001029	0.0000121	0.0011438	0.0000007	0.0000619	0.0000074	0.0000087	0.0000101	0.0002799	0.0046713	0.0001426	0.0000976	0.0000054	0.0005833	

в	0.0014149	0.0046235	0.0006755	0.0008383	0.0000161	0.0000955	0.0000444	0.0002590	0.0012474	0.0000094	0.0026993	0.0002166	0.0002785	0.0000047	0.0003263
г	0.0007899	0.0001171	0.0000276	0.0006856	0.0000034	0.0000034	0.0000000	0.0000511	0.0000081	0.0000000	0.0000000	0.0000000	0.0000094	0.0000040	0.0000054
д	0.0016961	0.0005174	0.0002644	0.0020406	0.0000040	0.0000706	0.0001978	0.0000727	0.0001157	0.0000007	0.0005968	0.0009325	0.0000235	0.0000081	0.0004515
е	0.0064408	0.0064105	0.0061171	0.0006694	0.0000706	0.0011666	0.0002032	0.0016208	0.0011243	0.0006984	0.0000000	0.0000000	0.0003485	0.0004467	0.0006203
ж	0.0000040	0.0001292	0.0000289	0.0002779	0.0000000	0.0000007	0.0000007	0.0000484	0.0000007	0.0000000	0.0000000	0.0000458	0.0000094	0.0000013	0.0000161
з	0.0002664	0.0001857	0.0000861	0.0003303	0.0000040	0.0000047	0.0000094	0.0000848	0.0000161	0.0000000	0.0004286	0.0001951	0.0000370	0.0000047	0.0005396
и	0.0010920	0.0044183	0.0053078	0.0005961	0.0000861	0.0016538	0.0009628	0.0027417	0.0005416	0.0001864	0.0000000	0.0000000	0.0002725	0.0002846	0.0022203
й	0.0003021	0.0013719	0.0005625	0.0001662	0.0000444	0.0000787	0.0000706	0.0003229	0.0001076	0.0000094	0.0000000	0.0000000	0.0000713	0.0000020	0.0000733
к	0.0016571	0.0015104	0.0006916	0.0016921	0.0000148	0.0000861	0.0000155	0.0001695	0.0000296	0.0000081	0.0000000	0.0000000	0.0001265	0.0000013	0.0001063
л	0.0001433	0.0022404	0.0003822	0.0015138	0.0000229	0.0000404	0.0000141	0.0006223	0.0000404	0.0000040	0.0010476	0.0039783	0.0001494	0.0011337	0.0020050
м	0.0002590	0.0009043	0.0003600	0.0032826	0.0000262	0.0000908	0.0000195	0.0004999	0.0000680	0.0000094	0.0010435	0.0000801	0.0000794	0.0000067	0.0005490
н	0.0002086	0.0018502	0.0007031	0.0029745	0.0000262	0.0000908	0.0004010	0.0004629	0.0000249	0.0001790	0.0032799	0.0013860	0.0000464	0.0001736	0.0019713
о	0.0067267	0.0090546	0.0088582	0.0007246	0.0002536	0.0006183	0.0001292	0.0031299	0.0014788	0.0001810	0.0000000	0.0000000	0.0005255	0.0008915	0.0010610
п	0.0069978	0.0000161	0.0000565	0.0005773	0.0000013	0.0000000	0.0000161	0.0000357	0.0000148	0.0000034	0.0002826	0.0000397	0.0000013	0.0000034	0.0003774
р	0.0000323	0.0003176	0.0005389	0.0025674	0.0000175	0.0001090	0.0000323	0.0000895	0.0002382	0.0000357	0.0013793	0.0009668	0.0000040	0.0001265	0.0010018
с	0.0002207	0.0009251	0.0121616	0.0007152	0.0000175	0.0001749	0.0000706	0.0005288	0.0000794	0.0000040	0.0003707	0.0031824	0.0000619	0.0001413	0.0033997
т	0.0028520	0.0014385	0.0003351	0.0014115	0.0000498	0.0000498	0.0000740	0.0005019	0.0000141	0.0000188	0.0016632	0.0070483	0.0001386	0.0000538	0.0006775
у	0.0007589	0.0017782	0.0015360	0.0001258	0.0000585	0.0003929	0.0000114	0.0014600	0.0007993	0.0002261	0.0000000	0.0000000	0.0000848	0.0010799	0.0002362
ф	0.0002826	0.0000161	0.0000229	0.0000289	0.0000081	0.0000000	0.0000000	0.0000020	0.0000000	0.0000000	0.0000054	0.0000599	0.0000000	0.0000013	0.0000013
х	0.0001662	0.0003593	0.0001063	0.0002617	0.0000114	0.0000215	0.0000108	0.0000585	0.0000235	0.0000020	0.0000000	0.0000000	0.0000007	0.0000296	0.0000020
ц	0.0000067	0.0000283	0.0000128	0.0002442	0.0000000	0.0000007	0.0000094	0.0000061	0.0000013	0.0000000	0.0001689	0.0000000	0.0000081	0.0000000	0.0000054
ч	0.0000854	0.0000976	0.0049976	0.0010590	0.0000000	0.0000087	0.0000007	0.0000161	0.0001904	0.0000000	0.0000000	0.0001796	0.0000040	0.0000000	0.0000202
ш	0.0000040	0.0000047	0.0000767	0.0003189	0.0000013	0.0000034	0.0000027	0.0000027	0.0000020	0.0000000	0.0000000	0.0005430	0.0000020	0.0000040	0.0000007
щ	0.0000061	0.0000013	0.0000000	0.0000996	0.0000000	0.0000000	0.0000000	0.0000000	0.0000000	0.0000000	0.0000000	0.0000262	0.0000007	0.0000000	0.0000000
ы	0.0003902	0.0013369	0.0009971	0.0001002	0.0000121	0.0009984	0.0000074	0.0003014	0.0004858	0.0000034	0.0000000	0.0000000	0.0000498	0.0000007	0.0000545
ь	0.0002523	0.0021240	0.0006311	0.0002711	0.0000336	0.0001056	0.0001285	0.0006405	0.0004992	0.0000209	0.0000000	0.0000000	0.0002893	0.0004427	0.0007811
э	0.0000081	0.0000047	0.0031952	0.0000000	0.0000020	0.0000000	0.0000000	0.0000000	0.0000007	0.0000000	0.0000000	0.0000000	0.0000000	0.0000000	0.0000000
ю	0.0001843	0.0005349	0.0004481	0.0001171	0.0000316	0.0000370	0.0000249	0.0002281	0.0000632	0.0003781	0.0000000	0.0000000	0.0000592	0.0000397	0.0000525
я	0.0003162	0.0021617	0.0020938	0.0002543	0.0000390	0.0003229	0.0000619	0.0005396	0.0000942	0.0002328	0.0000000	0.0000000	0.0001763	0.0001137	0.0002792

H2_ns = 3.9491384

R2_ns = 0.2028700

Таблиця частот біграм з пробілом, що не перетинаються, ентропія та надлишковість.
Від А до О, від Р до Я та пробіл.

	а	б	в	г	д	е	ж	з	и	й	к	л	м	н	о	п
а	0.0000221	0.0005655	0.0029299	0.0005615	0.0020738	0.0009791	0.0010681	0.0039702	0.0001303	0.0005014	0.0039872	0.0091541	0.0022925	0.0044286	0.0000391	0.0005728
б	0.0007932	0.0000011	0.0000816	0.0000051	0.0000306	0.0018947	0.0000051	0.0000023	0.0007372	0.0000000	0.0002266	0.0007621	0.0000544	0.0002850	0.0018738	0.0000000
в	0.0052015	0.0000159	0.0000238	0.0000300	0.0002963	0.0040648	0.0000000	0.0005060	0.0043005	0.0000000	0.0001456	0.0007406	0.0001371	0.0012052	0.0066712	0.0001960
г	0.0007955	0.0000000	0.0000074	0.0000028	0.0010885	0.0003394	0.0000045	0.0000006	0.0007134	0.0000000	0.0001128	0.0014256	0.0000000	0.0001643	0.0079659	0.0000011
д	0.0040133	0.0000210	0.0007230	0.0000040	0.0000368	0.0042773	0.0000130	0.0000011	0.0019418	0.0000000	0.0001977	0.0007553	0.0000657	0.0014465	0.0035084	0.0000629
е	0.0000606	0.0012256	0.0024018	0.0034750	0.0023282	0.0020517	0.0007496	0.0011077	0.0001054	0.0023763	0.0016794	0.0054740	0.0038070	0.0069653	0.0002130	0.0012947
ж	0.0013157	0.0000414	0.0000000	0.0000051	0.0007304	0.0038546	0.0000108	0.0000000	0.0013876	0.0000000	0.0000850	0.0000034	0.0000023	0.0009406	0.0000419	0.0000000
з	0.0050655	0.0001513	0.0007644	0.0004901	0.0007281	0.0002249	0.0000992	0.0000555	0.0002867	0.0000000	0.0000895	0.0001745	0.0002629	0.0018460	0.0003405	0.0000204
и	0.0000975	0.0003666	0.0025486	0.0003972	0.0014749	0.0022948	0.0002969	0.0017021	0.0005893	0.0012681	0.0015128	0.0042093	0.0026478	0.0039294	0.0001088	0.0001462
й	0.0000057	0.0000068	0.0000164	0.0000000	0.0001977	0.0000028	0.0000000	0.0000011	0.0000062	0.0000000	0.0000674	0.0000895	0.0000533	0.0002714	0.0000125	0.0000091
к	0.0073404	0.0000000	0.0001632	0.0000000	0.0000057	0.0004759	0.0000136	0.0000040	0.0027180	0.0000000	0.0000181	0.0003128	0.0000011	0.0005910	0.0074469	0.0000023
л	0.0072407	0.0000164	0.0000108	0.0001105	0.0000142	0.0049810	0.0003570	0.0000034	0.0050156	0.0000000	0.0002731	0.0002436	0.0000040	0.0002861	0.0055074	0.0000442
м	0.0023577	0.0000136	0.0000193	0.0000289	0.0000045	0.0028166	0.0000000	0.0000040	0.0025084	0.0000000	0.0000555	0.0000924	0.0000431	0.0012137	0.0029968	0.0000601
н	0.0104397	0.0000130	0.0000147	0.0000980	0.0006040	0.0093960	0.0000040	0.0000136	0.0073693	0.0000000	0.0002448	0.0000045	0.0000028	0.0027724	0.0092550	0.0000051
о	0.0000181	0.0032722	0.0069959	0.0043884	0.0040008	0.0023611	0.0018007	0.0009695	0.0006918	0.0029135	0.0014907	0.0052292	0.0045980	0.0079489	0.0002045	0.0010913
п	0.0010919	0.0000068	0.0000000	0.0000000	0.0000000	0.0019916	0.0000000	0.0000000	0.0007706	0.0000000	0.0000555	0.0004346	0.0000000	0.0000612	0.0077631	0.0000227
р	0.0065392	0.0002147	0.0002895	0.0003400	0.0003207	0.0049992	0.0002844	0.0000147	0.0041878	0.0000000	0.0005360	0.0000487	0.0001258	0.0006323	0.0065653	0.0000567

c	0.0016930	0.0000493	0.0017282	0.0000102	0.0002992	0.0041685	0.0000244	0.0000176	0.0012573	0.0000000	0.0048037	0.0024047	0.0007763	0.0008998	0.0022914	0.0015916
t	0.0047193	0.0000147	0.0025373	0.0000062	0.0001145	0.0047555	0.0000045	0.0000000	0.0036342	0.0000000	0.0004584	0.0002527	0.0000215	0.0010250	0.0136909	0.0000499
y	0.0000436	0.0004969	0.0010692	0.0009224	0.0014454	0.0002436	0.0016205	0.0002816	0.0000125	0.0001128	0.0007241	0.0010754	0.0011083	0.0002153	0.0000255	0.0005496
ф	0.0000844	0.0000000	0.0000000	0.0000000	0.0000000	0.0001162	0.0000000	0.0000000	0.0002952	0.0000000	0.0000000	0.0000232	0.0000062	0.0000017	0.0000929	0.0000000
x	0.0007768	0.0000017	0.0000935	0.0000000	0.0000034	0.0000334	0.0000000	0.0000000	0.0001082	0.0000000	0.0000040	0.0000833	0.0000465	0.0000963	0.0021055	0.0000068
ц	0.0004448	0.0000028	0.0000448	0.0000000	0.0000028	0.0006374	0.0000000	0.0000017	0.0001847	0.0000000	0.0001099	0.0000000	0.0000051	0.0000062	0.0002691	0.0000096
ч	0.0021197	0.0000000	0.0000057	0.0000006	0.0000000	0.0032008	0.0000000	0.0000006	0.0012278	0.0000000	0.0002958	0.0000419	0.0000045	0.0004646	0.0000465	0.0000000
ш	0.0008561	0.0000000	0.0000793	0.0000000	0.0000000	0.0021401	0.0000000	0.0000000	0.0014386	0.0000000	0.0003360	0.0005116	0.0000045	0.0003473	0.0002635	0.0000057
щ	0.0002748	0.0000000	0.0000000	0.0000000	0.0000000	0.0012414	0.0000000	0.0000000	0.0006312	0.0000000	0.0000000	0.0000000	0.0000000	0.0000300	0.0000011	0.0000000
ы	0.0000017	0.0004403	0.0009349	0.0000850	0.0001309	0.0009020	0.0000238	0.0000589	0.0000102	0.0011304	0.0001575	0.0024415	0.0010998	0.0003173	0.0000045	0.0001105
ь	0.0000130	0.0000884	0.0000414	0.0000510	0.0000317	0.0005762	0.0000011	0.0001603	0.0003955	0.0000000	0.0011343	0.0000136	0.0002091	0.0011321	0.0000312	0.0000130
э	0.0000000	0.0000011	0.0000000	0.0000023	0.0000017	0.0000000	0.0000000	0.0000057	0.0000000	0.0000028	0.0000374	0.0000402	0.0000017	0.0000227	0.0000000	0.0000204
ю	0.0000074	0.0005961	0.0000085	0.0000051	0.0002283	0.0000000	0.0000068	0.0000040	0.0000023	0.0000000	0.0000091	0.0000176	0.0000153	0.0000312	0.0000068	0.0000108
я	0.0000040	0.0000249	0.0002646	0.0002635	0.0005111	0.0000731	0.0002720	0.0002986	0.0000436	0.0001133	0.0001094	0.0007757	0.0003071	0.0006408	0.0000159	0.0000584
	0.0029197	0.0063568	0.0139176	0.0032710	0.0068854	0.0049691	0.0024143	0.0039141	0.0104885	0.0000023	0.0085002	0.0033441	0.0052604	0.0153879	0.0126654	0.0133062
	р	с	т	у	ф	х	ц	ч	ш	щ	ы	ь	э	ю	я	
a	0.0022353	0.0033634	0.0039759	0.0000521	0.0002442	0.0008034	0.0001145	0.0007474	0.0006046	0.0001926	0.0000000	0.0000000	0.0000079	0.0007961	0.0022902	0.0176532
б	0.0010414	0.0000759	0.0000074	0.0009241	0.0000000	0.0000516	0.0000057	0.0000057	0.0000079	0.0002357	0.0039340	0.0001201	0.0000000	0.0000045	0.0004737	0.0003649
в	0.0010103	0.0032926	0.0002187	0.0006255	0.0000000	0.0000544	0.0000221	0.0000748	0.0010097	0.0000074	0.0022732	0.0001824	0.0000000	0.0000023	0.0002482	0.0051590
г	0.0006476	0.0000289	0.0000062	0.0005575	0.0000011	0.0000011	0.0000000	0.0000357	0.0000051	0.0000000	0.0000000	0.0000000	0.0000000	0.0000034	0.0000000	0.0006482
д	0.0014041	0.0003218	0.0001768	0.0017004	0.0000000	0.0000578	0.0001660	0.0000255	0.0000963	0.0000000	0.0005026	0.0007853	0.0000006	0.0000068	0.0003734	0.0009037
e	0.0049680	0.0036756	0.0043391	0.0000907	0.0000102	0.0007451	0.0001524	0.0008578	0.0008805	0.0005774	0.0000000	0.0000000	0.0000011	0.0003745	0.0003349	0.0177410
ж	0.0000006	0.0000901	0.0000000	0.0002295	0.0000000	0.0000000	0.0000006	0.0000380	0.0000000	0.0000000	0.0000000	0.0000385	0.0000000	0.0000011	0.0000000	0.0002601
з	0.0001926	0.0000499	0.0000125	0.0002465	0.0000000	0.0000000	0.0000040	0.0000521	0.0000062	0.0000000	0.0003609	0.0001643	0.0000000	0.0000040	0.0004499	0.0009581
и	0.0005116	0.0020483	0.0036320	0.0000119	0.0000130	0.0012420	0.0007740	0.0017865	0.0003717	0.0001405	0.0000000	0.0000000	0.0000034	0.0002363	0.0016483	0.0171801
й	0.0000045	0.0004788	0.0002431	0.0000006	0.0000040	0.0000000	0.0000397	0.0000686	0.0000493	0.0000000	0.0000000	0.0000000	0.0000011	0.0000000	0.0000006	0.0067897
к	0.0013123	0.0008816	0.0003932	0.0013372	0.0000011	0.0000000	0.0000068	0.0000000	0.0000102	0.0000000	0.0000000	0.0000000	0.0000000	0.0000000	0.0000000	0.0044212
л	0.0000017	0.0011508	0.0000249	0.0011275	0.0000040	0.0000000	0.0000006	0.0001105	0.0000034	0.0000000	0.0008822	0.0033503	0.0000000	0.0009547	0.0016466	0.0068140
м	0.0000708	0.0001156	0.0000538	0.0025962	0.0000011	0.0000000	0.0000091	0.0000204	0.0000040	0.0000045	0.0008788	0.0000674	0.0000040	0.0000028	0.0003672	0.0065534
н	0.0000822	0.0010097	0.0004012	0.0023061	0.0000153	0.0000011	0.0003292	0.0002425	0.0000011	0.0001502	0.0027622	0.0011672	0.0000017	0.0001462	0.0016143	0.0051391
о	0.0051765	0.0056037	0.0062089	0.0000561	0.0001553	0.0003247	0.0000850	0.0015208	0.0011723	0.0001400	0.0000000	0.0000000	0.0000453	0.0007479	0.0004742	0.0223390
п	0.0058927	0.0000079	0.0000459	0.0004861	0.0000011	0.0000000	0.0000136	0.0000295	0.0000125	0.0000028	0.0002380	0.0000334	0.0000006	0.0000028	0.0003173	0.0000357
р	0.0000164	0.0001972	0.0004357	0.0021457	0.0000102	0.0000907	0.0000266	0.0000640	0.0001989	0.0000289	0.0011615	0.0008142	0.0000000	0.0001060	0.0008357	0.0006091
с	0.0001150	0.0005779	0.0100074	0.0004890	0.0000074	0.0001275	0.0000550	0.0004074	0.0000578	0.0000017	0.0003116	0.0026801	0.0000000	0.0001179	0.0028472	0.0027514
т	0.0023061	0.0008142	0.0000793	0.0011009	0.0000204	0.0000096	0.0000572	0.0002601	0.0000000	0.0000119	0.0014007	0.0059352	0.0000034	0.0000442	0.0004963	0.0041385
у	0.0005252	0.0009610	0.0010006	0.0000057	0.0000334	0.0002788	0.0000034	0.0008233	0.0006578	0.0001864	0.0000000	0.0000000	0.0000147	0.0009088	0.0001094	0.0061488
ф	0.0002374	0.0000102	0.0000176	0.0000238	0.0000068	0.0000000	0.0000000	0.0000000	0.0000000	0.0000000	0.0000045	0.0000504	0.0000000	0.0000011	0.0000000	0.0000357
х	0.0000640	0.0000402	0.0000108	0.0001422	0.0000000	0.0000000	0.0000000	0.0000000	0.0000085	0.0000000	0.0000000	0.0000006	0.0000006	0.0000000	0.0000006	0.0025956
ц	0.0000023	0.0000062	0.0000051	0.0002011	0.0000000	0.0000000	0.0000074	0.0000017	0.0000011	0.0000000	0.0001422	0.0000000	0.0000057	0.0000000	0.0000017	0.0001654
ч	0.0000567	0.0000057	0.0041680	0.0008607	0.0000000	0.0000000	0.0000000	0.0000000	0.0001581	0.0000000	0.0000000	0.0001513	0.0000000	0.0000000	0.0000000	0.0007139
ш	0.0000011	0.0000000	0.0000601	0.0002680	0.0000000	0.0000000	0.0000023	0.0000000	0.0000017	0.0000000	0.0000000	0.0004573	0.0000000	0.0000034	0.0000000	0.0000601
щ	0.0000051	0.0000000	0.0000000	0.0000839	0.0000000	0.0000000	0.0000000	0.0000000	0.0000000	0.0000000	0.0000000	0.0000221	0.0000000	0.0000000	0.0000000	0.0000074
ы	0.0002623	0.0007298	0.0006754	0.0000062	0.0000006	0.0007859	0.0000000	0.0001677	0.0003989	0.0000006	0.0000000	0.0000000	0.0000000	0.0000000	0.0000023	0.0039736
ь	0.0000034	0.0007910	0.0000720	0.0000028	0.0000028	0.0000011	0.0000969	0.0000589	0.0003847	0.0000142	0.0000000	0.0000000	0.0000045	0.0003728	0.0004992	0.0098244
э	0.0000068	0.0000040	0.0026908	0.0000000	0.0000000	0.0000000	0.0000000	0.0000000	0.0000000	0.0000000	0.0000000	0.0000000	0.0000000	0.0000000	0.0000000	0.0000028
ю	0.0000380	0.0001643	0.0002470	0.0000017	0.0000017	0.0000102	0.0000102	0.0000408	0.0000238	0.0003139	0.0000000	0.0000000	0.0000000	0.0000323	0.0000000	0.0031583
я	0.0000453	0.0008091	0.0012987	0.0000051	0.0000000	0.0001790	0.0000329	0.0001218	0.0000453	0.0001904	0.0000000	0.0000000	0.0000000	0.0000941	0.0001275	0.0105276
	0.0036586	0.0152661	0.0074571	0.0039192	0.0004726	0.0014596	0.0002470	0.0059601	0.0006652	0.0000958	0.0000000	0.0000000	0.0027441	0.0000289	0.0020947	0.0000000
H3 = 3.8124724																
R3 = 0.2375055																

Таблиця частот біграм без пробіла, що не перетинаються, ентропія та надлишковість.
Від А до П, від Р до Я.

	а	б	в	г	д	е	ж	з	и	й	к	л	м	н	о	п
а	0.0004992	0.0014452	0.0053933	0.0010980	0.0033796	0.0020279	0.0014506	0.0051504	0.0015448	0.0005954	0.0060493	0.0113188	0.0033311	0.0073842	0.0019276	0.0023515
б	0.0009581	0.0000034	0.0001077	0.0000087	0.0000384	0.0022674	0.0000061	0.0000047	0.0009171	0.0000000	0.0002732	0.0009076	0.0000680	0.0003492	0.0023495	0.0000067
в	0.0062532	0.0001743	0.0003216	0.0003176	0.0007253	0.0050778	0.0000612	0.0006964	0.0053334	0.0000000	0.0006816	0.0009978	0.0003647	0.0018563	0.0083537	0.0008195
г	0.0009554	0.0000397	0.0000949	0.0000135	0.0013503	0.0004293	0.0000141	0.0000249	0.0008928	0.0000000	0.0001588	0.0017022	0.0000188	0.0002698	0.0095076	0.0001097
д	0.0047871	0.0000706	0.0009332	0.0000283	0.0000774	0.0051107	0.0000289	0.0000384	0.0023872	0.0000000	0.0003082	0.0009177	0.0001056	0.0018550	0.0042273	0.0001433
е	0.0003277	0.0024699	0.0047575	0.0045550	0.0037853	0.0027787	0.0011613	0.0020844	0.0015260	0.0028218	0.0028696	0.0070088	0.0057526	0.0099671	0.0017769	0.0034879
ж	0.0015657	0.0000619	0.0000209	0.0000074	0.0008827	0.0045913	0.0000141	0.0000094	0.0016679	0.0000000	0.0001144	0.0000081	0.0000108	0.0011566	0.0000727	0.0000316
з	0.0060264	0.0002039	0.0009978	0.0006163	0.0009345	0.0002987	0.0001305	0.0000881	0.0003822	0.0000000	0.0002079	0.0002308	0.0003835	0.0022721	0.0004582	0.0001325
и	0.0003445	0.0012851	0.0050932	0.0008955	0.0026368	0.0033056	0.0005699	0.0025809	0.0017608	0.0015071	0.0028763	0.0054048	0.0037590	0.0066037	0.0018718	0.0022035
й	0.0004017	0.0002866	0.0006789	0.0002092	0.0006863	0.0001191	0.0001770	0.0001702	0.0007542	0.0000000	0.0005598	0.0002738	0.0003378	0.0009117	0.0005362	0.0007139
к	0.0087991	0.0003734	0.0005450	0.0000895	0.0002227	0.0006782	0.0002348	0.0000962	0.0035236	0.0000000	0.0003075	0.0004757	0.0001931	0.0013611	0.0092681	0.0003263
л	0.0087271	0.0002072	0.0008054	0.0002792	0.0002436	0.0064342	0.0004629	0.0001366	0.0064241	0.0000000	0.0007852	0.0006049	0.0001561	0.0009379	0.0076035	0.0005665
м	0.0029281	0.0002900	0.0006755	0.0002503	0.0003586	0.0034664	0.0001359	0.0001837	0.0036009	0.0000000	0.0005295	0.0002779	0.0002590	0.0020958	0.0041903	0.0007488
н	0.0127324	0.0003371	0.0006123	0.0002267	0.0009426	0.0112542	0.0000653	0.0002220	0.0090985	0.0000000	0.0004885	0.0000928	0.0001352	0.0040181	0.0113256	0.0006997
о	0.0002839	0.0052399	0.0107530	0.0056880	0.0058152	0.0036406	0.0026973	0.0017594	0.0022190	0.0034596	0.0028776	0.0067208	0.0062653	0.0120583	0.0027727	0.0033903
п	0.0012979	0.0000101	0.0000040	0.0000000	0.0000000	0.0023663	0.0000020	0.0000013	0.0009211	0.0000000	0.0000693	0.0005167	0.0000013	0.0000747	0.0092217	0.0000276
р	0.0077812	0.0002833	0.0003976	0.0004212	0.0004084	0.0059504	0.0003411	0.0000357	0.0050583	0.0000000	0.0006977	0.0000659	0.0001669	0.0008222	0.0078592	0.0001292
с	0.0020777	0.0001615	0.0022802	0.0001009	0.0005235	0.0050179	0.0002160	0.0000834	0.0015926	0.0000000	0.0059565	0.0029274	0.0010422	0.0015011	0.0029025	0.0021093
т	0.0056961	0.0002833	0.0034603	0.0000774	0.0003452	0.0058481	0.0000949	0.0001023	0.0046451	0.0000000	0.0007845	0.0004003	0.0002039	0.0017083	0.0165581	0.0004212
у	0.0002106	0.0008007	0.0018617	0.0012138	0.0020407	0.0004346	0.0020259	0.0004838	0.0007118	0.0001346	0.0013403	0.0014075	0.0015643	0.0011613	0.0005073	0.0012070
ф	0.0001070	0.0000034	0.0000094	0.0000000	0.0000000	0.0001379	0.0000000	0.0000000	0.0003539	0.0000000	0.0000000	0.0000276	0.0000087	0.0000061	0.0001110	0.0000000
х	0.0009716	0.0001225	0.0003532	0.0000881	0.0001823	0.0001023	0.0000518	0.0000693	0.0003923	0.0000000	0.0001763	0.0002166	0.0001675	0.0003512	0.0027572	0.0003041
ц	0.0005342	0.0000121	0.0000733	0.0000081	0.0000114	0.0007623	0.0000027	0.0000074	0.0002362	0.0000000	0.0001426	0.0000013	0.0000101	0.0000269	0.0003351	0.0000309
ч	0.0025285	0.0000256	0.0001198	0.0000128	0.0000343	0.0038108	0.0000054	0.0000350	0.0015085	0.0000000	0.0003828	0.0000599	0.0000168	0.0006338	0.0001164	0.0001077
ш	0.0010180	0.0000061	0.0000989	0.0000000	0.0000000	0.0025412	0.0000007	0.0000013	0.0017137	0.0000000	0.0004030	0.0006089	0.0000114	0.0004158	0.0003162	0.0000155
щ	0.0003263	0.0000000	0.0000013	0.0000000	0.0000000	0.0014741	0.0000000	0.0000000	0.0007495	0.0000000	0.0000000	0.0000000	0.0000000	0.0000363	0.0000020	0.0000020
ы	0.0000720	0.0007569	0.0015286	0.0001992	0.0003768	0.0011869	0.0000861	0.0002436	0.0004097	0.0013423	0.0004205	0.0029786	0.0014661	0.0009655	0.0002927	0.0005739
ь	0.0001601	0.0004023	0.0011384	0.0002200	0.0004777	0.0013315	0.0001110	0.0004844	0.0013551	0.0000000	0.0020252	0.0002234	0.0005981	0.0024242	0.0009285	0.0009238
э	0.0000000	0.0000027	0.0000020	0.0000027	0.0000020	0.0000000	0.0000000	0.0000067	0.0000000	0.0000034	0.0000444	0.0000478	0.0000000	0.0000283	0.0000000	0.0000242
ю	0.0000976	0.0007946	0.0002577	0.0000908	0.0004266	0.0001050	0.0001265	0.0000908	0.0003324	0.0000000	0.0002422	0.0001130	0.0001426	0.0003055	0.0003129	0.0002920
я	0.0003256	0.0004777	0.0014075	0.0005658	0.0011014	0.0005187	0.0005046	0.0006654	0.0010503	0.0001352	0.0008289	0.0011734	0.0007260	0.0024719	0.0008121	0.0010368
	р	с	т	у	ф	х	ц	ч	ш	щ	ы	ь	э	ю	я	
а	0.0030721	0.0059652	0.0056685	0.0005652	0.0003411	0.0011041	0.0001601	0.0015515	0.0008000	0.0002368	0.0000000	0.0000000	0.0003923	0.0009514	0.0030391	
б	0.0012373	0.0001029	0.0000121	0.0011438	0.0000007	0.0000619	0.0000074	0.0000087	0.0000101	0.0002799	0.0046714	0.0001426	0.0000976	0.0000054	0.0005833	
в	0.0014149	0.0046236	0.0006755	0.0008383	0.0000161	0.0000955	0.0000444	0.0002590	0.0012474	0.0000094	0.0026994	0.0002166	0.0002785	0.0000047	0.0003263	
г	0.0007899	0.0001171	0.0000276	0.0006856	0.0000034	0.0000034	0.0000000	0.0000511	0.0000081	0.0000000	0.0000000	0.0000000	0.0000094	0.0000040	0.0000054	
д	0.0016962	0.0005174	0.0002644	0.0020407	0.0000040	0.0000706	0.0001978	0.0000727	0.0001157	0.0000000	0.0005968	0.0009325	0.0000235	0.0000081	0.0004515	
е	0.0064409	0.0064106	0.0061173	0.0006695	0.0000706	0.0011667	0.0002032	0.0016208	0.0011243	0.0006984	0.0000000	0.0000000	0.0003485	0.0004468	0.0006203	
ж	0.0000040	0.0001292	0.0000289	0.0002779	0.0000000	0.0000007	0.0000000	0.0000484	0.0000000	0.0000000	0.0000000	0.0000458	0.0000094	0.0000013	0.0000161	
з	0.0002664	0.0001857	0.0000861	0.0003304	0.0000040	0.0000047	0.0000094	0.0000848	0.0000161	0.0000000	0.0004286	0.0001951	0.0000370	0.0000047	0.0005396	
и	0.0010920	0.0044184	0.0053079	0.0005961	0.0000861	0.0016538	0.0009628	0.0027417	0.0005416	0.0001864	0.0000000	0.0000000	0.0002725	0.0002846	0.0022203	
й	0.0003021	0.0013719	0.0005625	0.0001662	0.0000444	0.0000787	0.0000706	0.0003230	0.0001077	0.0000094	0.0000000	0.0000000	0.0000713	0.0000020	0.0000733	
к	0.0016572	0.0015105	0.0006917	0.0016921	0.0000148	0.0000861	0.0000155	0.0001696	0.0000296	0.0000081	0.0000000	0.0000000	0.0001265	0.0000013	0.0001063	
л	0.0001433	0.0022405	0.0003822	0.0015138	0.0000229	0.0000404	0.0000141	0.0006224	0.0000404	0.0000040	0.0010476	0.0039784	0.0001494	0.0011337	0.0020050	
м	0.0002590	0.0009043	0.0003600	0.0032827	0.0000262	0.0000908	0.0000195	0.0004999	0.0000680	0.0000094	0.0010435	0.0000801	0.0000794	0.0000067	0.0005490	
н	0.0002086	0.0018503	0.0007031	0.0029745	0.0000262	0.0000908	0.0004010	0.0004629	0.0000249	0.0001790	0.0032800	0.0013860	0.0000464	0.0001736	0.0019714	
о	0.0067268	0.0090548	0.0088583	0.0007246	0.0002537	0.0006183	0.0001292	0.0031300	0.0014789	0.0001810	0.0000000	0.0000000	0.0005255	0.0008915	0.0010610	
п	0.0069980	0.0000161	0.0000565	0.0005773	0.0000000	0.0000000	0.0000161	0.0000357	0.0000148	0.0000034	0.0002826	0.0000397	0.0000013	0.0000034	0.0003775	
р	0.0000323	0.0003176	0.0005389	0.0025675	0.0000175	0.0001090	0.0000323	0.0000895	0.0002382	0.0000357	0.0013793	0.0009668	0.0000040	0.0001265	0.0010018	

H3_ns = 3.9490027
R3_ns = 0.2028974

20 символів

[illegible]

Нерівність для значення надлишковості: $0,592 < R < 0,7126$

30 символів

[illegible]

Нерівність для значення надлишковості: $0,5336 < R < 0,6726$

Код:

```
#include <Windows.h>
#include <iostream>
#include <iomanip>
#include <fstream>
#include <chrono>
#include <filesystem>
#include <string>
#include <boost/algorithm/string.hpp>
#include <boost/regex.hpp>

// russian alphabet
std::string alpha[32]
{
    "а", "б", "в", "г", "д", "е", "ж", "з", "и", "й", "к", "л", "м", "н", "о", "п",
    "р", "с", "т", "у", "ф", "х", "ц", "ч", "ш", "щ", "ы", "ь", "э", "ю", "я", " "
};
//
// functions
size_t find_pair(std::string, std::string);
std::wstring toUNICODE(void*, int64_t);
std::string to1251(std::wstring);
std::string clear_the_text(std::string);
int get_occurrences(std::string, std::string);

void get_monogramms_data(std::string, double**, double**, double*,
double*, double*, double*);
void get_crossed_bigramm(std::string, std::string, std::string*,
std::string*, double**, double**, double*, double*, double*, double*);
void get_ncrossed_bigramm(std::string, std::string, std::string*,
std::string*, double**, double**, double*, double*, double*, double*);
//

int main()
{
    // setting console output to 1251 (russian ascii)
    SetConsoleCP(1251);
    SetConsoleOutputCP(1251);
    //

    // getting file name to parse
    std::string filename = "";
    std::cout << "Введите имя файла для обработки (Только
UTF-8): ";
    std::cin >> filename;
    //

    // reading the file
    FILE* file_ptr = NULL;
    fopen_s(&file_ptr, filename.c_str(), "rb");
    int64_t file_size =
std::experimental::filesystem::file_size(filename);
    char* data_ptr = (char*)malloc(file_size + 1);
    ZeroMemory(data_ptr, file_size + 1);
    fread(data_ptr, 1, file_size, file_ptr);
    fclose(file_ptr);
    //

    auto counter_start =
std::chrono::high_resolution_clock::now();
    // converting to unicode
    std::wstring data_unicode = toUNICODE(data_ptr,
file_size);
    //

    // converting to 1251
    std::string data_1251 = to1251(data_unicode);
    //

std::string to1251(std::wstring unicode_data)
{
    int64_t mb_buffer_size = WideCharToMultiByte(1251, 0,
unicode_data.c_str(), unicode_data.length(), 0, 0, 0, 0) + 1;

    char* mb_data_ptr = (char*)malloc(mb_buffer_size);
    ZeroMemory(mb_data_ptr, mb_buffer_size);

    WideCharToMultiByte(1251, 0, unicode_data.c_str(),
unicode_data.length(), mb_data_ptr, mb_buffer_size, 0, 0);

    std::string mb_string(mb_data_ptr);

    return mb_string;
}

std::string clear_the_text(std::string data)
{
    std::string data_ptr;

    for (int i = 0; i < data.length(); i++)
    {
        if (data[i] >= 'а' && data[i] <= 'я' || data[i] >= 'А' &&
data[i] <= 'Я' || data[i] == ' ')
        {
            data_ptr += data[i];
        }
    }

    boost::to_lower(data_ptr, std::locale("1251"));
    boost::replace_all(data_ptr, "ё", "е");
    boost::replace_all(data_ptr, "ъ", "ъ");
    data_ptr = boost::regex_replace(data_ptr, boost::regex("\\s+"), "
");

    if (data_ptr[0] == ' ')
    {
        data_ptr.erase(0);
    }
    if (data_ptr[data_ptr.length() - 1] == ' ')
    {
        data_ptr.erase(data_ptr.length() - 1);
    }

    return data_ptr;
}

int get_occurrences(std::string data, std::string keyword)
{
    int counter = 0;
    int64_t position = 0;

    while ((position = data.find(keyword, position)) !=
std::string::npos)
    {
        counter++;
        position += keyword.length();
    }

    return counter;
}

void get_monogramms_data(std::string text, double** array_ptr, double**
array_ns_ptr, double* entropy, double* entropy_ns, double* nadlishok,
double* nadlishok_ns)
{
    // variables initialization
    double* freq = new double[32];
    double* freq_ns = new double[31];
```

```

// clearing the text
std::string data = clear_the_text(data_1251);
std::string data_ns = data;
boost::replace_all(data_ns, " ", "");
//
auto counter_end =
std::chrono::high_resolution_clock::now();
double counter =
double(std::chrono::duration_cast<std::chrono::milliseconds>(counter_end - counter_start).count()) / 1000;

// writing the cleared files
std::ofstream cleared("cleared.txt");
cleared << data;
cleared.close();
std::ofstream cleared_ns("cleared_ns.txt");
cleared_ns << data_ns;
cleared_ns.close();
//

// getting monograms
double* monogramm_array = NULL, *
monogramm_ns_array = NULL;
double entropy = 0, entropy_ns = 0;
double nadlishok = 0, nadlishok_ns = 0;
counter_start = std::chrono::high_resolution_clock::now();
get_monograms_data(data, &monogramm_array,
&monogramm_ns_array, &entropy, &entropy_ns, &nadlishok,
&nadlishok_ns);
counter_end = std::chrono::high_resolution_clock::now();
counter +=
double(std::chrono::duration_cast<std::chrono::milliseconds>(counter_end - counter_start).count()) / 1000;

std::cout << "Монограммы:\n";
for (int i = 0; i < 32; i++)
{
    std::cout << alpha[i] << ": " << std::fixed <<
std::setprecision(7) << monogramm_array[i] << std::endl;
}
std::cout << "H1 = " << std::fixed << std::setprecision(7) <<
entropy << std::endl;
std::cout << "R1 = " << std::fixed << std::setprecision(7) <<
nadlishok << std::endl << std::endl;

for (int i = 0; i < 31; i++)
{
    std::cout << alpha[i] << ": " << std::fixed <<
std::setprecision(7) << monogramm_ns_array[i] << std::endl;
}
std::cout << "H1_ns = " << std::fixed << std::setprecision(7)
<< entropy_ns << std::endl;
std::cout << "R1_ns = " << std::fixed << std::setprecision(7)
<< nadlishok_ns << std::endl << std::endl;
//

// getting bigramms
double* bigramm_crossed = NULL, * bigramm_ns_crossed
= NULL;
double bigramm_crossed_entropy = 0,
bigramm_crossed_ns_entropy = 0;
double bigramm_crossed_nadlishok = 0,
bigramm_crossed_ns_nadlishok = 0;
std::string pairs_crossed, ns_pairs_crossed;
counter_start = std::chrono::high_resolution_clock::now();
get_crossed_bigramm(data, data_ns, &pairs_crossed,
&ns_pairs_crossed, &bigramm_crossed, &bigramm_ns_crossed,
&bigramm_crossed_entropy, &bigramm_crossed_ns_entropy,
&bigramm_crossed_nadlishok, &bigramm_crossed_ns_nadlishok);
counter_end = std::chrono::high_resolution_clock::now();

```

```

double monogramm_entropy = 0, monogramm_entropy_ns = 0;
double monogramm_nadlishok = 0, monogramm_nadlishok_ns =
0;

int monogramm_size = 0;
//

// getting freq
for (int i = 0; i < 31; i++)
{
    freq[i] = get_occurrences(text, alpha[i]);
    freq_ns[i] = freq[i];
    monogramm_size += freq[i];
}
freq[31] = get_occurrences(text, alpha[31]);

for (int i = 0; i < 31; i++)
{
    freq_ns[i] /= monogramm_size;
}

monogramm_size += freq[31];

for (int i = 0; i < 32; i++)
{
    freq[i] /= monogramm_size;
}
//

// getting entropy
for (int i = 0; i < 31; i++)
{
    if (freq[i] != 0)
    {
        monogramm_entropy += freq[i] *
log2(freq[i]);
    }

    if (freq_ns[i] != 0)
    {
        monogramm_entropy_ns += freq_ns[i] *
log2(freq_ns[i]);
    }
}
if (freq[31] != 0)
{
    monogramm_entropy += freq[31] * log2(freq[31]);
}
monogramm_entropy *= (-1);
monogramm_entropy_ns *= (-1);
//

// getting nadlishok
monogramm_nadlishok = double(1 - monogramm_entropy /
log2(32));
monogramm_nadlishok_ns = double(1 -
monogramm_entropy_ns / log2(31));
//

// finalizing
*array_ptr = freq;
*array_ns_ptr = freq_ns;
*entropy = monogramm_entropy;
*entropy_ns = monogramm_entropy_ns;
*nadlishok = monogramm_nadlishok;
*nadlishok_ns = monogramm_nadlishok_ns;
//

}

void get_crossed_bigramm(std::string text, std::string text_ns, std::string*
pairs, std::string* pairs_ns, double** array_ptr, double** array_ns_ptr,
double* entropy, double* entropy_ns, double* nadlishok, double*
nadlishok_ns)

```

```

        counter +=
double(std::chrono::duration_cast<std::chrono::milliseconds>(counter_
end - counter_start).count()) / 1000;

        std::cout << "Пересекающиеся биграммы:\n";
        for (int i = 0; i < 32; i++)
        {
                std::cout << "\t\t" << alpha[i];
        }
        std::cout << std::endl << std::endl;

        for (int i = 0; i < 32; i++)
        {
                std::cout << alpha[i] << "\t";
                for (int j = 0; j < 32; j++)
                {
                        std::string pair = alpha[i] + alpha[j];
                        size_t position = find_pair(pairs_crossed,
pair);

                        if (position != std::string::npos)
                        {
                                std::cout << std::fixed <<
std::setprecision(7) << bigramm_crossed[position / 2] << "\t";
                        }
                        else
                        {
                                std::cout << std::fixed <<
std::setprecision(7) << double(0) << "\t";
                        }
                        std::cout << std::endl << std::endl;
                }
                std::cout << "H2 = " << std::fixed << std::setprecision(7) <<
bigramm_crossed_entropy << std::endl;
                std::cout << "R2 = " << std::fixed << std::setprecision(7) <<
bigramm_crossed_nadlishok << std::endl << std::endl;

                for (int i = 0; i < 31; i++)
                {
                        std::cout << "\t\t" << alpha[i];
                }
                std::cout << std::endl << std::endl;

                for (int i = 0; i < 31; i++)
                {
                        std::cout << alpha[i] << "\t";
                        for (int j = 0; j < 31; j++)
                        {
                                std::string pair = alpha[i] + alpha[j];
                                size_t position =
find_pair(ns_pairs_crossed, pair);
                                if (position != std::string::npos)
                                {
                                        std::cout << std::fixed <<
std::setprecision(7) << bigramm_ns_crossed[position / 2] << "\t";
                                }
                                else
                                {
                                        std::cout << std::fixed <<
std::setprecision(7) << double(0) << "\t";
                                }
                                std::cout << std::endl << std::endl;
                        }
                }
                std::cout << "H2_ns = " << std::fixed << std::setprecision(7)
<< bigramm_crossed_ns_entropy << std::endl;
                std::cout << "R2_ns = " << std::fixed << std::setprecision(7)
<< bigramm_crossed_ns_nadlishok << std::endl << std::endl;
                //

                // getting ncrossed bigramms

```

```

{
        // variables initialization
        std::string bigram_pairs, bigram_ns_pairs;
        double* freq = new double[32 * 32];
        double* freq_ns = new double[31 * 31];
        double bigram_entropy = 0, bigram_ns_entropy = 0;
        double bigram_nadlishok = 0, bigram_ns_nadlishok = 0;
        int bigram_size = 0, bigram_ns_size = 0;
        //

        // getting freq
        for (int i = 0, j = -1; i < text.length(); i++)
        {
                std::string pair = text.substr(i, 2);

                if (find_pair(bigram_pairs, pair) == std::string::npos)
                {
                        bigram_pairs += pair;
                        freq[++j] = get_occurrences(text, pair);
                        bigram_size += freq[j];
                }
        }
        for (int i = 0, j = -1; i < text_ns.length(); i++)
        {
                std::string pair = text_ns.substr(i, 2);

                if (find_pair(bigram_ns_pairs, pair) ==
std::string::npos)
                {
                        bigram_ns_pairs += pair;
                        freq_ns[++j] = get_occurrences(text_ns,
pair);

                        bigram_ns_size += freq_ns[j];
                }
        }
        for (int i = 0; i < bigram_pairs.length() / 2; i++)
        {
                freq[i] /= bigram_size;
        }
        for (int i = 0; i < bigram_ns_pairs.length() / 2; i++)
        {
                freq_ns[i] /= bigram_ns_size;
        }
        //

        // getting entropy
        for (int i = 0; i < bigram_pairs.length() / 2; i++)
        {
                if (freq[i] != 0)
                {
                        bigram_entropy += freq[i] * log2(freq[i]);
                }
        }
        for (int i = 0; i < bigram_ns_pairs.length() / 2; i++)
        {
                if (freq_ns[i] != 0)
                {
                        bigram_ns_entropy += freq_ns[i] *
log2(freq_ns[i]);
                }
        }
        bigram_entropy *= (-1);
        bigram_ns_entropy *= (-1);
        bigram_entropy /= 2;
        bigram_ns_entropy /= 2;
        //

        // getting nadlishok
        bigram_nadlishok = double(1 - bigram_entropy / log2(32));
        bigram_ns_nadlishok = double(1 - bigram_ns_entropy /
log2(31));
        //

```

```

double* bigramm_ncrossed = NULL, *
bigramm_ns_ncrossed = NULL;
double bigramm_ncrossed_entropy = 0,
bigramm_ncrossed_ns_entropy = 0;
double bigramm_ncrossed_nadlishok = 0,
bigram_ncrossed_ns_nadlishok = 0;
std::string pairs, ns_pairs;
counter_start = std::chrono::high_resolution_clock::now();
get_ncrossed_bigramm(data, data_ns, &pairs, &ns_pairs,
&bigramm_ncrossed, &bigramm_ns_ncrossed,
&bigramm_ncrossed_entropy, &bigramm_ncrossed_ns_entropy,
&bigramm_ncrossed_nadlishok, &bigram_ncrossed_ns_nadlishok);
counter_end = std::chrono::high_resolution_clock::now();
counter +=
double(std::chrono::duration_cast<std::chrono::milliseconds>(counter_
end - counter_start).count()) / 1000;

std::cout << "Не пересекающиеся биграммы:\n";
for (int i = 0; i < 32; i++)
{
    std::cout << "\t\t" << alpha[i];
}
std::cout << std::endl << std::endl;

for (int i = 0; i < 32; i++)
{
    std::cout << alpha[i] << "\t";
    for (int j = 0; j < 32; j++)
    {
        std::string pair = alpha[i] + alpha[j];
        size_t position = find_pair(pairs, pair);
        if (position != std::string::npos)
        {
            std::cout << std::fixed <<
std::setprecision(7) << bigramm_ncrossed[position / 2] << "\t";
        }
        else
        {
            std::cout << std::fixed <<
std::setprecision(7) << double(0) << "\t";
        }
    }
    std::cout << std::endl << std::endl;
}
std::cout << "H3 = " << std::fixed << std::setprecision(7) <<
bigramm_ncrossed_entropy << std::endl;
std::cout << "R3 = " << std::fixed << std::setprecision(7) <<
bigramm_ncrossed_nadlishok << std::endl << std::endl;

for (int i = 0; i < 31; i++)
{
    std::cout << "\t\t" << alpha[i];
}
std::cout << std::endl << std::endl;

for (int i = 0; i < 31; i++)
{
    std::cout << alpha[i] << "\t";
    for (int j = 0; j < 31; j++)
    {
        std::string pair = alpha[i] + alpha[j];
        size_t position = find_pair(ns_pairs,
pair);
        if (position != std::string::npos)
        {
            std::cout << std::fixed <<
std::setprecision(7) << bigramm_ns_ncrossed[position / 2] << "\t";
        }
        else
        {
            std::cout << std::fixed <<
std::setprecision(7) << double(0) << "\t";

```

```

// finalizing
*array_ptr = freq;
*array_ns_ptr = freq_ns;
*pairs = bigram_pairs;
*pairs_ns = bigram_ns_pairs;
*entropy = bigram_entropy;
*entropy_ns = bigram_ns_entropy;
*nadlishok = bigram_nadlishok;
*nadlishok_ns = bigram_ns_nadlishok;
//
}

void get_ncrossed_bigramm(std::string text, std::string text_ns, std::string*
pairs, std::string* pairs_ns, double** array_ptr, double** array_ns_ptr,
double* entropy, double* entropy_ns, double* nadlishok, double*
nadlishok_ns)
{
    // variables initialization
    std::string bigram_pairs, bigram_ns_pairs;
    double* freq = new double[32 * 32];
    double* freq_ns = new double[31 * 31];
    double bigram_entropy = 0, bigram_ns_entropy = 0;
    double bigram_nadlishok = 0, bigram_ns_nadlishok = 0;
    int bigram_size = 0, bigram_ns_size = 0;
    //

    // getting freq
    for (int i = 0, j = -1; i < text.length(); i+=2)
    {
        std::string pair = text.substr(i, 2);

        if (find_pair(bigram_pairs, pair) == std::string::npos)
        {
            bigram_pairs += pair;
            freq[++j] = get_occurrences(text, pair);
            bigram_size += freq[j];
        }
    }
    for (int i = 0, j = -1; i < text_ns.length(); i+=2)
    {
        std::string pair = text_ns.substr(i, 2);

        if (find_pair(bigram_ns_pairs, pair) ==
std::string::npos)
        {
            bigram_ns_pairs += pair;
            freq_ns[++j] = get_occurrences(text_ns,
pair);
            bigram_ns_size += freq_ns[j];
        }
    }
    for (int i = 0; i < bigram_pairs.length() / 2; i++)
    {
        freq[i] /= bigram_size;
    }
    for (int i = 0; i < bigram_ns_pairs.length() / 2; i++)
    {
        freq_ns[i] /= bigram_ns_size;
    }
    //

    // getting entropy
    for (int i = 0; i < bigram_pairs.length() / 2; i++)
    {
        if (freq[i] != 0)
        {
            bigram_entropy += freq[i] * log2(freq[i]);
        }
    }
    for (int i = 0; i < bigram_ns_pairs.length() / 2; i++)
    {

```

<pre> } } std::cout << std::endl << std::endl; } std::cout << "H3_ns = " << std::fixed << std::setprecision(7) << bigramm_nscrossed_ns_entropy << std::endl; std::cout << "R3_ns = " << std::fixed << std::setprecision(7) << bigram_ncrossed_ns_nadlishok << std::endl << std::endl; // // overall execution time std::cout << "Программа обработала файл за: " << counter << " секунд.\n"; // system("pause"); return 0; } size_t find_pair(std::string pairs_data, std::string pair) { for (int i = 0; i < pairs_data.length(); i += 2) { std::string pair_to_cmp = pairs_data.substr(i, 2); if (pair == pair_to_cmp) { return i; } } return std::string::npos; } std::wstring toUNICODE(void* buffer, int64_t buffer_size) { int64_t unicode_buffer_size = MultiByteToWideChar(CP_UTF8, 0, (LPCCH)buffer, buffer_size, 0, 0) * sizeof(wchar_t) + 2; wchar_t* unicode_data_ptr = (wchar_t*)malloc(unicode_buffer_size); ZeroMemory(unicode_data_ptr, unicode_buffer_size); MultiByteToWideChar(CP_UTF8, 0, (LPCCH)buffer, buffer_size, unicode_data_ptr, unicode_buffer_size); std::wstring unicode_wstring(unicode_data_ptr); return unicode_wstring; } </pre>	<pre> if (freq_ns[i] != 0) { bigram_ns_entropy += freq_ns[i] * log2(freq_ns[i]); } } bigram_entropy *= (-1); bigram_ns_entropy *= (-1); bigram_entropy /= 2; bigram_ns_entropy /= 2; // // getting nadlishok bigram_nadlishok = double(1 - bigram_entropy / log2(32)); bigram_ns_nadlishok = double(1 - bigram_ns_entropy / log2(31)); // // finalyizing *array_ptr = freq; *array_ns_ptr = freq_ns; *pairs = bigram_pairs; *pairs_ns = bigram_ns_pairs; *entropy = bigram_entropy; *entropy_ns = bigram_ns_entropy; *nadlishok = bigram_nadlishok; *nadlishok_ns = bigram_ns_nadlishok; // } </pre>
---	--

Висновки:

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