

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ УКРАЇНИ «КИЇВСЬКИЙ
ПОЛІТЕХНІЧНИЙ ІНСТИТУТ ім. І. СІКОРСЬКОГО»

Комп'ютерний практикум №1

Експериментальна оцінка ентропії на символ джерела відкритого тексту

Виконала:

Студентка групи ФБ-83

Захаряш Ксенія

Перевірив:

Чорний О.М.

Київ

2020

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 видалено 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.

Хід роботи

Під час виконання роботи виникали труднощі з підбором найкращих методів та структур даних для обробки тексту та коректним виведенням таблиць. Структури даних були використані найрізноманітніші (словники, масиви, і т.д.), а для виведення таблиць дуже добре підійшли дата фрейми.

Результати виконання роботи представлені у протоколі у вигляді таблиць та скріншотів виконання CoolPinkProgram.

Таблиця 1.0 Частоти монограм в тексті з пробілами

- 0.138536	Р 0.040921	У 0.018268	Х 0.009594
О 0.092314	В 0.038016	Ь 0.015965	Ж 0.007767
Е 0.076489	Л 0.034145	Б 0.015833	Ю 0.006467
И 0.072535	М 0.027429	Я 0.015660	Ш 0.005660
А 0.060701	К 0.026809	Г 0.013811	Ц 0.005457
Н 0.060164	Д 0.025043	З 0.012234	Щ 0.004019
Т 0.058827	П 0.023291	Ч 0.012020	Э 0.003938
С 0.046210	Ы 0.019101	Й 0.009885	Ф 0.002888

Таблиця 2.0 Частоти монограм в тексті без пробілів

О 0.107160	В 0.044130	Ь 0.018532	Ж 0.009016
Е 0.088790	Л 0.039637	Б 0.018379	Ю 0.007507
И 0.084200	М 0.031840	Я 0.018178	Ш 0.006570
А 0.070463	К 0.031121	Г 0.016032	Ц 0.006335
Н 0.069840	Д 0.029070	З 0.014201	Щ 0.004665
Т 0.068288	П 0.027036	Ч 0.013953	Э 0.004572
С 0.053642	Ы 0.022173	Й 0.011475	Ф 0.003353
Р 0.047501	У 0.021205	Х 0.011137	

Таблиця 1.1. Частоти біграм з перетинами в тексті з пробілами

	-	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
-	0,000000	0,002191	0,006468	0,013384	0,002921	0,007051	0,002384	0,001059	0,003206	0,009690	0,000004	0,007828	0,002093	0,005283	0,012184	0,007908	0,015048	0,005090	0,013665	0,005917	0,002970	0,001251	0,000845	0,000570	0,004124	0,000770	0,000017	0,000000	0,000000	0,003745	0,000115	0,000756
А	0,010647	0,000004	0,001980	0,003220	0,000424	0,001351	0,001529	0,000987	0,002020	0,000222	0,000428	0,003120	0,005831	0,003060	0,006184	0,000038	0,001093	0,002489	0,003076	0,005908	0,000092	0,000109	0,001003	0,000964	0,000885	0,000255	0,000419	0,000000	0,000000	0,000015	0,001545	0,001803
Б	0,000099	0,000601	0,000009	0,000017	0,000000	0,000001	0,001593	0,000006	0,000001	0,000639	0,000000	0,000026	0,000882	0,000038	0,000409	0,000000	0,000927	0,000301	0,000000	0,001308	0,000000	0,000179	0,000000	0,000000	0,000010	0,000695	0,003425	0,000168	0,000010	0,000099	0,000200	
В	0,008138	0,005828	0,000001	0,000170	0,000005	0,000115	0,004277	0,000000	0,000119	0,002755	0,000000	0,000289	0,001285	0,000152	0,001574	0,006025	0,000190	0,001210	0,001500	0,000323	0,000635	0,000022	0,000036	0,000017	0,000009	0,000506	0,000022	0,002498	0,000051	0,000004	0,000000	0,000261
Г	0,000955	0,001081	0,000000	0,000010	0,000233	0,000962	0,000256	0,000000	0,000000	0,001440	0,000000	0,000059	0,000786	0,000008	0,000132	0,006009	0,000001	0,001299	0,000013	0,000005	0,000517	0,000000	0,000001	0,000000	0,000035	0,000004	0,000000	0,000000	0,000000	0,000004	0,000000	0,000000
Д	0,000699	0,003429	0,000004	0,000656	0,000015	0,000131	0,005218	0,000242	0,000003	0,002706	0,000000	0,000067	0,001052	0,000022	0,001582	0,004294	0,000259	0,000686	0,000771	0,000035	0,001482	0,000000	0,000019	0,000305	0,000029	0,000074	0,000000	0,000663	0,000261	0,000001	0,000003	0,000336
Е	0,017867	0,000342	0,000771	0,001478	0,001883	0,003662	0,001349	0,000986	0,001174	0,000163	0,002451	0,001592	0,004836	0,004554	0,010814	0,000571	0,000599	0,005865	0,006718	0,005340	0,000155	0,000242	0,000722	0,000168	0,000772	0,000601	0,000450	0,000000	0,000000	0,000026	0,000170	0,000169
Ж	0,000067	0,000695	0,000023	0,000000	0,000001	0,001062	0,002992	0,000000	0,000000	0,001059	0,000000	0,000067	0,000001	0,000012	0,001591	0,000019	0,000000	0,000000	0,000008	0,000000	0,000058	0,000000	0,000000	0,000001	0,000097	0,000000	0,000000	0,000000	0,000014	0,000000	0,000000	0,000000
З	0,001055	0,003317	0,000213	0,000930	0,000093	0,000562	0,000269	0,000018	0,000036	0,000564	0,000001	0,000211	0,000182	0,001155	0,001199	0,001022	0,000005	0,000532	0,000038	0,000001	0,000259	0,000000	0,000000	0,000014	0,000001	0,000010	0,000000	0,000345	0,000023	0,000000	0,000000	0,000178
И	0,017073	0,000704	0,000553	0,002295	0,000597	0,001418	0,004270	0,000605	0,003441	0,002193	0,001621	0,003156	0,002644	0,003634	0,004113	0,000946	0,000143	0,002691	0,002996	0,005793	0,000013	0,000410	0,002856	0,001162	0,002372	0,000434	0,000088	0,000000	0,000000	0,000001	0,000660	0,003654
Й	0,008009	0,000004	0,000003	0,000015	0,000008	0,000111	0,000005	0,000000	0,000005	0,000000	0,000000	0,000041	0,000017	0,000076	0,000364	0,000023	0,000004	0,000005	0,000776	0,000159	0,000000	0,000003	0,000000	0,000059	0,000058	0,000141	0,000000	0,000000	0,000000	0,000001	0,000000	0,000000
К	0,002610	0,005644	0,000000	0,000469	0,000000	0,000003	0,000567	0,000133	0,000008	0,003503	0,000000	0,000015	0,000461	0,000001	0,000128	0,008662	0,000001	0,001063	0,000426	0,001460	0,001515	0,000001	0,000003	0,000124	0,000000	0,000006	0,000000	0,000000	0,000000	0,000001	0,000003	0,000000
Л	0,000945	0,004177	0,000004	0,000001	0,000205	0,000019	0,004622	0,000581	0,000003	0,007879	0,000000	0,000114	0,000790	0,000008	0,000196	0,003603	0,000003	0,000000	0,000186	0,000028	0,001045	0,000004	0,000005	0,000000	0,000005	0,000000	0,000000	0,000506	0,006137	0,000000	0,001090	0,001989
М	0,006853	0,002066	0,000049	0,000032	0,000008	0,000000	0,005183	0,000000	0,000000	0,004487	0,000000	0,000045	0,000197	0,000497	0,000692	0,003161	0,000521	0,000006	0,000106	0,000020	0,001595	0,000003	0,000000	0,000019	0,000000	0,000000	0,000004	0,001347	0,000131	0,000005	0,000004	0,000398
Н	0,001300	0,007979	0,000005	0,000428	0,000168	0,000710	0,007886	0,000004	0,000013	0,011335	0,000000	0,000571	0,000006	0,000001	0,004038	0,011454	0,000010	0,000017	0,001910	0,001475	0,001401	0,000333	0,000004	0,000788	0,000100	0,000003	0,000038	0,006424	0,000923	0,000005	0,000068	0,000767
О	0,015339	0,000014	0,004825	0,007942	0,005441	0,004547	0,002220	0,001655	0,001337	0,001516	0,004255	0,001976	0,006447	0,005310	0,005629	0,000342	0,001456	0,004625	0,006807	0,005699	0,000022	0,000232	0,000665	0,000812	0,001734	0,000441	0,000264	0,000000	0,000000	0,000076	0,000229	0,000457
П	0,000165	0,001279	0,000000	0,000000	0,000000	0,000000	0,002368	0,000000	0,000000	0,001023	0,000000	0,000038	0,001037	0,000000	0,000149	0,008096	0,000174	0,007910	0,000029	0,000051	0,000434	0,000000	0,000000	0,000031	0,000000	0,000001	0,000000	0,000286	0,000038	0,000018	0,000000	0,000161
Р	0,000771	0,010121	0,000055	0,000345	0,000570	0,000186	0,007700	0,000415	0,000010	0,004658	0,000000	0,000097	0,000038	0,000670	0,000523	0,008284	0,000088	0,000046	0,000628	0,000590	0,002177	0,000008	0,000132	0,000029	0,000033	0,000095	0,000001	0,001655	0,000415	0,000004	0,000023	0,000553
С	0,001814	0,000849	0,000254	0,001178	0,000004	0,000247	0,002371	0,000004	0,000000	0,002184	0,000000	0,004232	0,002537	0,000421	0,000838	0,004188	0,002017	0,000802	0,000836	0,015173	0,000889	0,000026	0,000400	0,000024	0,000334	0,000177	0,000000	0,000170	0,000961	0,000005	0,000044	0,003233
Т	0,007495	0,006788	0,000020	0,004017	0,000000	0,000101	0,005710	0,000000	0,000000	0,005703	0,000000	0,000622	0,000081	0,000038	0,001682	0,010914	0,000024	0,004013	0,002071	0,000059	0,000970	0,000006	0,000005	0,000013	0,000078	0,000000	0,000003	0,001500	0,006447	0,000003	0,000024	0,000441
У	0,003555	0,000169	0,000300	0,000297	0,000854	0,001921	0,000377	0,000894	0,000119	0,000029	0,000019	0,000438	0,000496	0,000439	0,000551	0,000013	0,000946	0,000777	0,001007	0,001264	0,000001	0,000000	0,000137	0,000017	0,001007	0,000211	0,000786	0,000000	0,000000	0,000010	0,001597	0,000035
Ф	0,000026	0,000187	0,000000	0,000000	0,000005	0,000000	0,000273	0,000000	0,000000	0,001304	0,000000	0,000000	0,000209	0,000001	0,000003	0,000300	0,000000	0,000204	0,000040	0,000068	0,000141	0,000122	0,000000	0,000000	0,000000	0,000000	0,000000	0,000000	0,000008	0,000000	0,000000	0,000000
Х	0,006526	0,000143	0,000001	0,000050	0,000013	0,000004	0,000024	0,000000	0,000000	0,000079	0,000000	0,000000	0,000012	0,000004	0,000348	0,001998	0,000006	0,000190	0,000074	0,000009	0,000109	0,000001	0,000000	0,000000	0,000000	0,000000	0,000000	0,000000	0,000001	0,000001	0,000000	0,000000
Ц	0,000101	0,000565	0,000000	0,000018	0,000000	0,000000	0,001381	0,000000	0,000000	0,002909	0,000000	0,000020	0,000000	0,000000	0,000000	0,000095	0,000000	0,000000	0,000000	0,000001	0,000092	0,000000	0,000000	0,000000	0,000000	0,000000	0,000275	0,000000	0,000000	0,000000	0,000000	0,000000
Ч	0,000026	0,001625	0,000000	0,000004	0,000000	0,000000	0,004493	0,000000	0,000000	0,001794	0,000000	0,000104	0,000040	0,000000	0,001021	0,000005	0,000000	0,000072	0,000000	0,002482	0,000132	0,000000	0,000000	0,000000	0,000000	0,000170	0,000000	0,000000	0,000053	0,000000	0,000000	0,000000
Ш	0,000029	0,000511	0,000000	0,000026	0,000000	0,000000	0,001762	0,000000	0,000000	0,001355	0,000000	0,000170	0,000434	0,000005	0,000307	0,000149	0,000000	0,000005	0,000003	0,000500	0,000152	0,000000	0,000000	0,000000	0,000000	0,000000	0,000000	0,000000	0,000251	0,000000	0,000000	0,000000
Щ	0,000006	0,000380	0,000000	0,000000	0,000000	0,000000	0,002158	0,000000	0,000000	0,001221	0,000000	0,000000	0,000000	0,000000	0,000040	0,000000	0,000000	0,000022	0,000000	0,000000	0,000102	0,000000	0,000000	0,000000	0,000000	0,000000	0,000000	0,000000	0,000000	0,000090	0,000000	0,000000
Ы	0,006316	0,000000	0,000092	0,000710	0,000076	0,000042	0,002635	0,000065	0,000085	0,000055	0,001007	0,000055	0,001373	0,001369	0,000612	0,000003	0,000101	0,000245	0,000701	0,000616	0,000001	0,000000	0,0002325	0,000000	0,000117	0,000497	0,000000	0,000000	0,000000	0,000000	0,000000	0,000006
Ь	0,008046	0,000004	0,000046	0,000009	0,000236	0,000040	0,000284	0,000001	0,000313	0,000064	0,000000	0,000622	0,000000	0,00014																		

Таблиця 1.2. Частоти біграм без перетинів в тексті з пробілами

	-	А	Б	В	Г	Д	Є	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	
-	0,000000	0,002182	0,006478	0,013423	0,002961	0,007085	0,002377	0,001050	0,003245	0,009816	0,000008	0,007928	0,002124	0,005256	0,012178	0,007961	0,014757	0,005154	0,013712	0,005848	0,002948	0,001204	0,000815	0,000620	0,004260	0,000771	0,000018	0,000000	0,000000	0,003724	0,000113	0,000771	
А	0,010659	0,000003	0,002011	0,003322	0,000459	0,001388	0,001609	0,001022	0,001985	0,000207	0,000430	0,003225	0,005802	0,003130	0,006186	0,000046	0,001053	0,002503	0,003094	0,005989	0,000095	0,000108	0,000984	0,001025	0,000876	0,000256	0,000420	0,000000	0,000000	0,000005	0,001568	0,001770	
Б	0,000100	0,000620	0,000013	0,000020	0,000000	0,000000	0,001570	0,000008	0,000000	0,000617	0,000000	0,000026	0,000879	0,000038	0,000423	0,004142	0,000000	0,000953	0,000307	0,000000	0,001324	0,000000	0,000187	0,000000	0,000000	0,000008	0,000745	0,003415	0,000161	0,000010	0,000100	0,000218	
В	0,008125	0,005817	0,000000	0,000149	0,000010	0,000108	0,004293	0,000000	0,000123	0,002813	0,000000	0,000264	0,001224	0,000149	0,001598	0,005892	0,000215	0,001247	0,001488	0,000341	0,000658	0,000026	0,000038	0,000026	0,000010	0,000464	0,000023	0,002449	0,000059	0,000003	0,000000	0,000256	
Г	0,000984	0,001104	0,000000	0,000005	0,000225	0,000986	0,000279	0,000000	0,000000	0,001378	0,000000	0,000044	0,000725	0,000013	0,000126	0,006058	0,000000	0,001271	0,000013	0,000010	0,000530	0,000000	0,000000	0,000000	0,000033	0,000000	0,000000	0,000000	0,000000	0,000005	0,000000	0,000000	
Д	0,000676	0,003456	0,000003	0,000646	0,000015	0,000151	0,005141	0,000231	0,000000	0,002633	0,000000	0,000077	0,001032	0,000028	0,001547	0,004285	0,000220	0,000692	0,000786	0,000033	0,001445	0,000000	0,000018	0,000287	0,000036	0,000074	0,000000	0,000674	0,000282	0,000003	0,000005	0,000328	
Е	0,017862	0,000305	0,000825	0,001452	0,001898	0,003835	0,001365	0,001002	0,001245	0,000156	0,002574	0,001506	0,004875	0,004577	0,010697	0,000599	0,000569	0,005871	0,006660	0,005302	0,000159	0,000248	0,000740	0,000184	0,000766	0,000599	0,000451	0,000000	0,000000	0,000015	0,000169	0,000190	
Ж	0,000056	0,000674	0,000018	0,000000	0,000000	0,001063	0,003030	0,000000	0,000000	0,001030	0,000000	0,000061	0,000003	0,000010	0,001588	0,000020	0,000000	0,000000	0,000008	0,000000	0,000054	0,000000	0,000000	0,000003	0,000082	0,000000	0,000000	0,000000	0,000015	0,000000	0,000000	0,000000	
З	0,001040	0,003325	0,000218	0,000989	0,000100	0,000556	0,000300	0,000018	0,000033	0,000517	0,000000	0,000184	0,000154	0,001224	0,001132	0,001012	0,000005	0,000512	0,000051	0,000000	0,000251	0,000000	0,000015	0,000003	0,000018	0,000000	0,000330	0,000018	0,000000	0,000000	0,000200		
И	0,017024	0,000710	0,000558	0,002305	0,000617	0,001401	0,004288	0,000617	0,003394	0,002234	0,001534	0,003151	0,002554	0,003707	0,004180	0,000973	0,000141	0,002708	0,003074	0,005766	0,000015	0,000425	0,002792	0,001158	0,002328	0,000415	0,000079	0,000000	0,000000	0,000003	0,000630	0,003584	
Й	0,007954	0,000005	0,000000	0,000013	0,000005	0,000118	0,000003	0,000000	0,000005	0,000000	0,000000	0,000038	0,000018	0,000072	0,000333	0,000015	0,000003	0,000000	0,000720	0,000172	0,000000	0,000005	0,000000	0,000069	0,000054	0,000143	0,000000	0,000000	0,000000	0,000000	0,000000	0,000000	
К	0,002508	0,005587	0,000000	0,000482	0,000000	0,000000	0,000574	0,000100	0,000008	0,003520	0,000000	0,000018	0,000476	0,000000	0,000136	0,008804	0,000000	0,000991	0,000441	0,001486	0,001493	0,000003	0,000003	0,000115	0,000000	0,000008	0,000000	0,000000	0,000000	0,000003	0,000003	0,000000	
Л	0,000953	0,004134	0,000005	0,000003	0,000202	0,000023	0,004577	0,000561	0,000005	0,008048	0,000000	0,000092	0,000781	0,000005	0,000195	0,003730	0,000003	0,000000	0,000202	0,000031	0,001025	0,000003	0,000008	0,000000	0,000008	0,000000	0,000000	0,000000	0,000533	0,006025	0,000000	0,001078	0,001954
М	0,006870	0,002016	0,000056	0,000033	0,000008	0,000000	0,005113	0,000000	0,006401	0,000000	0,000056	0,000207	0,000484	0,000686	0,003125	0,000528	0,000008	0,000110	0,000013	0,001565	0,000005	0,000000	0,000020	0,000000	0,000000	0,000000	0,000005	0,001324	0,000143	0,000008	0,000003	0,000389	
Н	0,001291	0,008092	0,000003	0,000415	0,000167	0,000722	0,007700	0,000003	0,000018	0,011591	0,000000	0,000569	0,000008	0,000003	0,004029	0,011619	0,000008	0,000018	0,001957	0,001483	0,001376	0,000348	0,000005	0,000730	0,000082	0,000003	0,000046	0,006440	0,000981	0,000008	0,000064	0,000745	
О	0,015016	0,000010	0,004667	0,007949	0,005359	0,004521	0,002157	0,001701	0,001342	0,001560	0,004375	0,001990	0,006614	0,005226	0,005594	0,000348	0,001488	0,004642	0,006829	0,005712	0,000015	0,000231	0,000663	0,000784	0,001703	0,000412	0,000274	0,000000	0,000000	0,000072	0,000228	0,000384	
П	0,000141	0,001294	0,000000	0,000000	0,000000	0,000000	0,002377	0,000000	0,000000	0,001017	0,000000	0,000033	0,001055	0,000000	0,000159	0,008205	0,000164	0,008305	0,000031	0,000056	0,000448	0,000000	0,000000	0,000038	0,000000	0,000000	0,000000	0,000300	0,000044	0,000020	0,000000	0,000146	
Р	0,000781	0,009885	0,000049	0,000313	0,000579	0,000200	0,007598	0,000420	0,000015	0,004452	0,000000	0,000090	0,000044	0,000684	0,000517	0,008197	0,000095	0,000038	0,000605	0,000615	0,002234	0,000008	0,000141	0,000023	0,000028	0,000079	0,000003	0,001614	0,000438	0,000003	0,000028	0,000553	
С	0,001906	0,000835	0,000243	0,001178	0,000003	0,000254	0,002377	0,000000	0,000000	0,002167	0,000000	0,004191	0,002526	0,000430	0,000914	0,004147	0,001893	0,000776	0,000820	0,015085	0,000856	0,000023	0,000371	0,000023	0,000338	0,000205	0,000000	0,000169	0,001012	0,000003	0,000049	0,003181	
Т	0,007469	0,006624	0,000023	0,004011	0,000000	0,000105	0,005638	0,000000	0,000000	0,005781	0,000000	0,000630	0,000097	0,000046	0,001570	0,011074	0,000031	0,004068	0,002095	0,000049	0,000966	0,000008	0,000008	0,000013	0,000085	0,000000	0,000000	0,001465	0,006470	0,000003	0,000023	0,000418	
У	0,003550	0,000182	0,000305	0,000295	0,000858	0,001875	0,000387	0,000899	0,000105	0,000026	0,000008	0,000433	0,000469	0,000418	0,000528	0,000010	0,000927	0,000807	0,001066	0,001319	0,000003	0,000000	0,000143	0,000018	0,000104	0,000220	0,000779	0,000000	0,000000	0,000013	0,001621	0,000028	
Ф	0,000018	0,000195	0,000000	0,000000	0,000005	0,000000	0,000269	0,000000	0,000000	0,0001283	0,000000	0,000000	0,000210	0,000000	0,000300	0,000000	0,000223	0,000033	0,000079	0,000141	0,000128	0,000000	0,000000	0,000000	0,000000	0,000000	0,000000	0,000000	0,000008	0,000000	0,000000	0,000000	
Х	0,006727	0,000149	0,000000	0,000059	0,000010	0,000003	0,000026	0,000000	0,000000	0,000072	0,000000	0,000000	0,000015	0,000005	0,000330	0,001978	0,000005	0,000202	0,000079	0,000008	0,000108	0,000000	0,000000	0,000000	0,000000	0,000000	0,000000	0,000000	0,000000	0,000000	0,000000	0,000000	
Ц	0,000087	0,000584	0,000000	0,000023	0,000000	0,000000	0,001378	0,000000	0,000000	0,002848	0,000000	0,000023	0,000000	0,000000	0,000072	0,000000	0,000000	0,000000	0,000000	0,000003	0,000115	0,000000	0,000000	0,000000	0,000000	0,000000	0,000000	0,000000	0,000000	0,000000	0,000000	0,000000	
Ч	0,000023	0,001522	0,000000	0,000008	0,000000	0,000000	0,004565	0,000000	0,000000	0,001875	0,000000	0,000118	0,000041	0,000000	0,000989	0,000008	0,000000	0,000079	0,000000	0,002410	0,000120	0,000000	0,000000	0,000000	0,000000	0,000164	0,000000	0,000000	0,000059	0,000000	0,000000	0,000000	
Ш	0,000028	0,000507	0,000000	0,000023	0,000000	0,000000	0,001849	0,000000	0,000000	0,001350	0,000000	0,000169	0,000418	0,000005	0,000295	0,000133	0,000000	0,000003	0,000005	0,000523	0,000151	0,000000	0,000000	0,000000	0,000000	0,000000	0,000000	0,000000	0,000254	0,000000	0,000000	0,000000	
Щ	0,000008	0,000356	0,000000	0,000000	0,000000	0,000000	0,002170	0,000000	0,000000	0,001209	0,000000	0,000000	0,000000	0,000000	0,000033	0,000000	0,000000	0,000023	0,000000	0,000000	0,000133	0,000000	0,000000	0,000000	0,000000	0,000000	0,000000	0,000000	0,000000	0,000085	0,000000	0,000000	
Ы	0,006337	0,000000	0,000105	0,000722	0,000074	0,000051	0,002672	0,000072	0,000092	0,000049	0,000991	0,000041	0,001363	0,001452	0,000622	0,000003	0,000085	0,000259	0,000684	0,000625	0,000003	0,000000	0,002236	0,000000	0,000120	0,000502	0,000000	0,000000	0,000000	0,000000	0,000000	0,000003	
Ь	0,007890	0,000000	0,000051	0,000010	0,000213	0,000041	0,000272	0,000003	0,000313	0,000051	0,000000	0,000669	0,0000																				

Таблиця 2.1. Частоти біграм з перетинами в тексті без пробілів

	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	
А	0,000229	0,002921	0,004920	0,000766	0,002266	0,001998	0,001197	0,002652	0,001206	0,000497	0,004303	0,006903	0,003971	0,008492	0,000673	0,002660	0,003290	0,004716	0,007339	0,000358	0,000207	0,001237	0,001175	0,001308	0,000314	0,000488	0,000000	0,000000	0,000404	0,001808	0,002166	
Б	0,000702	0,000013	0,000027	0,000000	0,000006	0,001852	0,000007	0,000003	0,000761	0,000000	0,000031	0,001024	0,000045	0,000479	0,004874	0,000004	0,001079	0,000358	0,000003	0,001530	0,000000	0,000208	0,000000	0,000000	0,000012	0,000807	0,003976	0,000195	0,000036	0,000114	0,000232	
В	0,007003	0,000339	0,000714	0,000471	0,000740	0,005162	0,000034	0,000280	0,003781	0,000000	0,001060	0,001592	0,000529	0,002602	0,007437	0,001012	0,001698	0,002863	0,001148	0,000864	0,000079	0,000117	0,000067	0,000178	0,000627	0,000025	0,002899	0,000059	0,000323	0,000025	0,000400	
Г	0,001282	0,000065	0,000117	0,000282	0,001194	0,000308	0,000003	0,000021	0,001756	0,000000	0,000141	0,000923	0,000042	0,000280	0,007026	0,000110	0,001543	0,000114	0,000052	0,000626	0,000012	0,000006	0,000003	0,000065	0,000004	0,000000	0,000000	0,000000	0,000042	0,000000	0,000013	
Д	0,003996	0,000024	0,000850	0,000027	0,000175	0,006069	0,000284	0,000031	0,003194	0,000000	0,000172	0,001274	0,000048	0,001942	0,005022	0,000376	0,000813	0,000954	0,000073	0,001739	0,000004	0,000025	0,000357	0,000048	0,000089	0,000000	0,000770	0,000303	0,000015	0,000003	0,000391	
Е	0,000644	0,001864	0,003799	0,002687	0,005300	0,001912	0,001305	0,001918	0,001597	0,002847	0,002853	0,005896	0,006258	0,014016	0,001876	0,003003	0,007649	0,009798	0,006974	0,000654	0,000526	0,001085	0,000335	0,001518	0,000944	0,000529	0,000000	0,000000	0,000525	0,000204	0,000274	
Ж	0,000807	0,000030	0,000009	0,000003	0,001237	0,003475	0,000000	0,000001	0,001230	0,000000	0,000079	0,000003	0,000018	0,001860	0,000027	0,000006	0,000006	0,000018	0,000001	0,000068	0,000000	0,000003	0,000003	0,000114	0,000000	0,000000	0,000000	0,000000	0,000016	0,000001	0,000000	0,000000
З	0,003879	0,000290	0,001166	0,000149	0,000717	0,000321	0,000031	0,000054	0,000721	0,000001	0,000335	0,000226	0,001374	0,001555	0,001258	0,000112	0,000653	0,000167	0,000071	0,000323	0,000013	0,000001	0,000019	0,000018	0,000016	0,000000	0,000400	0,000027	0,000095	0,000001	0,000210	
И	0,001137	0,001751	0,004741	0,001033	0,002688	0,005276	0,000806	0,004532	0,003953	0,001882	0,004811	0,003417	0,004848	0,006650	0,002212	0,002385	0,003754	0,005211	0,007544	0,000464	0,000642	0,003439	0,001409	0,003289	0,000563	0,000106	0,000000	0,000000	0,000489	0,000791	0,004375	
Й	0,000174	0,000284	0,000853	0,000149	0,000613	0,000152	0,000101	0,000202	0,000737	0,000000	0,000644	0,000091	0,000581	0,000931	0,000381	0,001183	0,000483	0,002035	0,000586	0,000183	0,000126	0,000045	0,000122	0,000245	0,000195	0,000001	0,000000	0,000000	0,000340	0,000009	0,000028	
К	0,006622	0,000122	0,000788	0,000088	0,000140	0,000697	0,000238	0,000073	0,004237	0,000000	0,000251	0,000592	0,000153	0,000346	0,010245	0,000282	0,001372	0,000778	0,001811	0,001830	0,000028	0,000018	0,000149	0,000094	0,000016	0,000001	0,000000	0,000000	0,000123	0,000010	0,000015	
Л	0,004863	0,000098	0,000126	0,000256	0,000079	0,005391	0,000682	0,000018	0,009209	0,000000	0,000180	0,000937	0,000046	0,000306	0,004243	0,000101	0,000033	0,000335	0,000089	0,001237	0,000016	0,000012	0,000007	0,000049	0,000001	0,000000	0,000587	0,007124	0,000030	0,001267	0,002313	
М	0,002516	0,000395	0,000755	0,000187	0,000445	0,006173	0,000061	0,000150	0,005791	0,000000	0,000520	0,000332	0,000974	0,001439	0,004242	0,001406	0,000290	0,000883	0,000280	0,002065	0,000062	0,000048	0,000054	0,000317	0,000039	0,000006	0,001564	0,000152	0,000183	0,000012	0,000501	
Н	0,009289	0,000126	0,000648	0,000230	0,000889	0,009172	0,000010	0,000064	0,013278	0,000001	0,000721	0,000010	0,000049	0,004856	0,013396	0,000199	0,000062	0,002367	0,001756	0,001661	0,000397	0,000016	0,000916	0,000147	0,000004	0,000045	0,007458	0,001072	0,000024	0,000082	0,000894	
О	0,000198	0,006760	0,011110	0,006534	0,006146	0,002975	0,002090	0,001913	0,002698	0,004940	0,003137	0,007796	0,006965	0,008070	0,001775	0,003537	0,006056	0,009493	0,007356	0,000452	0,000406	0,000877	0,000981	0,002642	0,000559	0,000306	0,000000	0,000000	0,000474	0,000271	0,000638	
П	0,001491	0,000001	0,000022	0,000006	0,000004	0,002753	0,000000	0,000004	0,001206	0,000000	0,000052	0,001206	0,000000	0,000222	0,009404	0,000213	0,009197	0,000052	0,000070	0,000505	0,000000	0,000000	0,000036	0,000006	0,000001	0,000000	0,000332	0,000045	0,000021	0,000000	0,000187	
Р	0,011765	0,000112	0,000498	0,000694	0,000259	0,008949	0,000482	0,000025	0,005477	0,000000	0,000161	0,000056	0,000801	0,000697	0,009671	0,000201	0,000083	0,000792	0,000717	0,002544	0,000019	0,000162	0,000039	0,000070	0,000114	0,000001	0,001921	0,000482	0,000031	0,000028	0,000650	
С	0,001041	0,000394	0,001502	0,000042	0,000461	0,002795	0,000010	0,000030	0,002670	0,000000	0,004994	0,002969	0,000549	0,001210	0,004963	0,002597	0,001024	0,001097	0,017767	0,001084	0,000045	0,000473	0,000051	0,000416	0,000207	0,000000	0,000198	0,001115	0,000086	0,000051	0,003803	
Т	0,007979	0,000516	0,005574	0,000135	0,000540	0,006780	0,000051	0,000201	0,007208	0,000000	0,001136	0,000263	0,000271	0,002722	0,013231	0,000993	0,004952	0,003247	0,000498	0,001344	0,000085	0,000033	0,000036	0,000464	0,000012	0,000003	0,001741	0,007484	0,000220	0,000030	0,000541	
У	0,000278	0,000456	0,000687	0,001066	0,002346	0,000494	0,001093	0,000222	0,000329	0,000022	0,000757	0,000644	0,000651	0,001146	0,000251	0,001574	0,001038	0,001525	0,001653	0,000061	0,000036	0,000187	0,000028	0,001461	0,000248	0,000913	0,000000	0,000000	0,000112	0,001857	0,000070	
Ф	0,000217	0,000001	0,000000	0,000006	0,000001	0,000317	0,000000	0,000000	0,001522	0,000000	0,000003	0,000242	0,000003	0,000348	0,000004	0,000238	0,000051	0,000079	0,000165	0,000141	0,000000	0,000000	0,000000	0,000000	0,000001	0,000000	0,000009	0,000000	0,000000	0,000000	0,000000	
Х	0,000305	0,000210	0,000581	0,000349	0,000373	0,000103	0,000056	0,000184	0,000648	0,000000	0,000479	0,000195	0,000346	0,000889	0,002661	0,000891	0,000607	0,000902	0,000281	0,000287	0,000117	0,000021	0,000064	0,000126	0,000296	0,000000	0,000000	0,000001	0,000135	0,000001	0,000027	
Ц	0,000660	0,000009	0,000024	0,000001	0,000006	0,001607	0,000000	0,000003	0,003382	0,000000	0,000037	0,000001	0,000001	0,000013	0,000117	0,000010	0,000001	0,000019	0,000003	0,000109	0,000003	0,000000	0,000000	0,000003	0,000000	0,000000	0,000320	0,000000	0,000003	0,000000	0,000000	
Ч	0,001887	0,000000	0,000010	0,000000	0,000000	0,005216	0,000001	0,000001	0,002086	0,000000	0,000122	0,000049	0,000000	0,001185	0,000006	0,000001	0,000083	0,000001	0,002883	0,000156	0,000000	0,000000	0,000000	0,000004	0,000198	0,000000	0,000000	0,000061	0,000001	0,000000	0,000000	
Ш	0,000596	0,000001	0,000034	0,000000	0,000003	0,002047	0,000000	0,000000	0,001574	0,000000	0,000198	0,000504	0,000009	0,000358	0,000174	0,000004	0,000006	0,000007	0,000580	0,000177	0,000001	0,000001	0,000000	0,000000	0,000000	0,000000	0,000000	0,000291	0,000001	0,000000	0,000000	
Щ	0,000442	0,000000	0,000000	0,000000	0,000000	0,000001	0,002505	0,000000	0,000000	0,001418	0,000000	0,000001	0,000000	0,000046	0,000000	0,000000	0,000025	0,000001	0,000000	0,000120	0,000000	0,000000	0,000000	0,000000	0,000000	0,000000	0,000000	0,000104	0,000000	0,000000	0,000000	
Ы	0,000125	0,000476	0,001576	0,000180	0,000385	0,003228	0,000106	0,000254	0,000703	0,001169	0,000464	0,001679	0,001802	0,001560	0,000456	0,001036	0,000480	0,001524	0,000962	0,000201	0,000039	0,002739	0,000004	0,000300	0,000587	0,000001	0,000000	0,000000	0,000100	0,000003	0,000036	
Ь	0,000132	0,000453	0,001100	0,000406	0,000479	0,000485	0,000128	0,000602	0,000746	0,000000	0,001161	0,000226	0,000500	0,004361	0,000456	0,000840	0,000293	0,002481	0,000584	0,000177	0,000051	0,000043	0,000051	0,000355	0,001470	0,000016	0,000000	0,000000	0,000248	0,000430	0,000257	
Э	0,000003	0,000006	0,000016	0,000015	0,000003	0,000000	0,000000	0,000000	0,000001	0,000006	0,001380	0,000141	0,000025	0,000080	0,000000	0,000077	0,000045	0,000051	0,002580	0,000000	0,000135	0,000003	0,000001	0,000000	0,000003	0,000000	0,000000	0,000000	0,000000	0,000000	0,000000	
Ю	0,000048	0,000223	0,000226	0,000080	0,000837	0,000028	0,000098	0,000199	0																							

Таблиця 2.2. Частоти біграм без перетинів в тексті без пробілів

	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
А	0,000217	0,002902	0,005070	0,000699	0,002412	0,001980	0,001207	0,002858	0,001201	0,000467	0,004294	0,007032	0,003946	0,008724	0,000702	0,002587	0,003411	0,004627	0,007229	0,000348	0,000193	0,001270	0,001207	0,001359	0,000327	0,000500	0,000000	0,000000	0,000428	0,001716	0,002209
Б	0,000687	0,000018	0,000042	0,000000	0,000006	0,001847	0,000003	0,000006	0,000752	0,000000	0,000036	0,000960	0,000056	0,000452	0,004927	0,000000	0,001172	0,000378	0,000006	0,001490	0,000000	0,000169	0,000000	0,000000	0,000015	0,000827	0,003976	0,000205	0,000030	0,000125	0,000250
В	0,006869	0,000342	0,000714	0,000458	0,000740	0,005108	0,000036	0,000274	0,003753	0,000000	0,001056	0,001576	0,000503	0,002617	0,007306	0,001073	0,001704	0,002866	0,001097	0,000868	0,000074	0,000110	0,000065	0,000149	0,000547	0,000030	0,002902	0,000062	0,000339	0,000024	0,000401
Г	0,001317	0,000074	0,000146	0,000268	0,001183	0,000324	0,000006	0,000021	0,001808	0,000000	0,000155	0,000880	0,000042	0,000288	0,006884	0,000095	0,001552	0,000140	0,000054	0,000586	0,000015	0,000003	0,000006	0,000071	0,000003	0,000000	0,000000	0,000000	0,000056	0,000000	0,000021
Д	0,003993	0,000021	0,000844	0,000045	0,000164	0,006128	0,000312	0,000036	0,003155	0,000000	0,000164	0,001192	0,000048	0,001957	0,005094	0,000387	0,000836	0,000972	0,000077	0,001662	0,000003	0,000015	0,000339	0,000045	0,000068	0,000000	0,000839	0,000306	0,000009	0,000003	0,000372
Е	0,000636	0,001957	0,003761	0,002700	0,005317	0,001894	0,001323	0,001909	0,001606	0,002748	0,002777	0,005893	0,006232	0,014157	0,001894	0,003015	0,007582	0,009836	0,007026	0,000633	0,000505	0,001059	0,000306	0,001508	0,001008	0,000538	0,000000	0,000000	0,000473	0,000202	0,000297
Ж	0,000761	0,000018	0,000012	0,000000	0,001252	0,003396	0,000000	0,000000	0,001234	0,000000	0,000080	0,000003	0,000012	0,001790	0,000021	0,000006	0,000000	0,000021	0,000003	0,000074	0,000000	0,000000	0,000000	0,000110	0,000000	0,000000	0,000000	0,000018	0,000000	0,000000	0,000000
З	0,003788	0,000294	0,001219	0,000158	0,000681	0,000339	0,000033	0,000059	0,000723	0,000000	0,000342	0,000184	0,001466	0,001558	0,001225	0,000107	0,000684	0,000155	0,000048	0,000333	0,000012	0,000003	0,000024	0,000003	0,000015	0,000000	0,000381	0,000024	0,000104	0,000000	0,000238
И	0,001100	0,001766	0,004769	0,001029	0,002658	0,005245	0,000841	0,004448	0,004029	0,001948	0,004835	0,003428	0,004894	0,006753	0,002174	0,002432	0,003741	0,005049	0,007556	0,000455	0,000613	0,003402	0,001406	0,003387	0,000592	0,000104	0,000000	0,000000	0,000446	0,000770	0,004416
Й	0,000161	0,000333	0,000874	0,000140	0,000621	0,000169	0,000101	0,000208	0,000785	0,000000	0,000645	0,0000104	0,000562	0,000990	0,000357	0,001163	0,000452	0,002141	0,000544	0,000155	0,000155	0,000042	0,000125	0,000247	0,000202	0,000000	0,000000	0,000000	0,000330	0,000018	0,000036
К	0,006444	0,000110	0,000740	0,000104	0,000122	0,000672	0,000232	0,000056	0,004276	0,000000	0,000271	0,000607	0,000161	0,000315	0,010455	0,000294	0,001406	0,000818	0,001805	0,001799	0,000027	0,000012	0,000167	0,000098	0,000018	0,000000	0,000000	0,000000	0,000134	0,000012	0,000006
Л	0,004909	0,000083	0,000128	0,000274	0,000054	0,005653	0,000663	0,000021	0,009224	0,000000	0,000169	0,000960	0,000045	0,000321	0,004145	0,000095	0,000024	0,000327	0,000101	0,001204	0,000018	0,000015	0,000009	0,000054	0,000000	0,000000	0,000595	0,007023	0,000030	0,001252	0,002310
М	0,002563	0,000407	0,000743	0,000190	0,000422	0,006307	0,000071	0,000167	0,005763	0,000000	0,000514	0,000342	0,001029	0,001475	0,004142	0,001389	0,000277	0,000892	0,000285	0,002052	0,000065	0,000033	0,000056	0,000285	0,000042	0,000003	0,001514	0,000158	0,000205	0,000009	0,000488
Н	0,009233	0,000107	0,000583	0,000238	0,000877	0,009001	0,000015	0,000065	0,013181	0,000000	0,000740	0,000012	0,000045	0,004752	0,013461	0,000202	0,000062	0,002325	0,001763	0,001591	0,000393	0,000012	0,000925	0,000143	0,000003	0,000045	0,007547	0,001065	0,000024	0,000071	0,000880
О	0,000205	0,006527	0,011246	0,006673	0,006194	0,002926	0,002233	0,001826	0,002724	0,004799	0,003155	0,007737	0,007006	0,007969	0,001858	0,003693	0,006096	0,009471	0,007157	0,000461	0,000395	0,000859	0,001002	0,002658	0,000547	0,000315	0,000000	0,000000	0,000494	0,000288	0,000657
П	0,001502	0,000003	0,000018	0,000003	0,000006	0,002724	0,000000	0,000006	0,001228	0,000000	0,000054	0,001151	0,000000	0,000226	0,009209	0,000199	0,009280	0,000062	0,000062	0,000550	0,000000	0,000000	0,000036	0,000006	0,000003	0,000000	0,000297	0,000039	0,000024	0,000000	0,000202
Р	0,011534	0,000125	0,000505	0,000669	0,000232	0,008944	0,000446	0,000018	0,005546	0,000000	0,000143	0,000051	0,000764	0,000684	0,009521	0,000217	0,000071	0,000779	0,000731	0,002563	0,000021	0,000143	0,000051	0,000062	0,000149	0,000003	0,001939	0,000556	0,000036	0,000036	0,000657
С	0,001017	0,000419	0,001567	0,000042	0,000428	0,002795	0,000015	0,000027	0,002587	0,000000	0,005016	0,002962	0,000529	0,001273	0,005055	0,002691	0,001044	0,001106	0,017713	0,001112	0,000048	0,000449	0,000054	0,000395	0,000214	0,000000	0,000205	0,001088	0,000101	0,000042	0,003791
Т	0,008055	0,000482	0,005718	0,000116	0,000479	0,006872	0,000048	0,000202	0,007116	0,000000	0,001189	0,000297	0,000285	0,002715	0,013440	0,000984	0,004992	0,003226	0,000544	0,001338	0,000101	0,000036	0,000039	0,000494	0,000018	0,000003	0,001731	0,007472	0,000217	0,000018	0,000541
У	0,000291	0,000446	0,000654	0,001050	0,002429	0,000482	0,001097	0,000247	0,000390	0,000024	0,000746	0,000705	0,000565	0,001210	0,000280	0,001626	0,001035	0,001481	0,001698	0,000071	0,000027	0,000181	0,000024	0,001514	0,000238	0,000913	0,000000	0,000000	0,000101	0,001861	0,000065
Ф	0,000220	0,000000	0,000000	0,000006	0,000003	0,000330	0,000000	0,000000	0,001558	0,000000	0,000006	0,000265	0,000003	0,000003	0,000318	0,000003	0,000235	0,000074	0,000077	0,000149	0,000131	0,000000	0,000000	0,000000	0,000000	0,000000	0,000006	0,000000	0,000000	0,000000	0,000000
Х	0,000303	0,000208	0,000574	0,000345	0,000381	0,000101	0,000056	0,000164	0,000696	0,000000	0,000503	0,000208	0,000333	0,000892	0,002706	0,000919	0,000621	0,000874	0,000312	0,000282	0,000122	0,000021	0,000065	0,000128	0,000300	0,000000	0,000000	0,000003	0,000122	0,000000	0,000042
Ц	0,000690	0,000006	0,000018	0,000000	0,000006	0,001612	0,000000	0,000006	0,003327	0,000000	0,000036	0,000000	0,000000	0,000015	0,000095	0,000009	0,000003	0,000024	0,000003	0,000095	0,000003	0,000000	0,000000	0,000003	0,000000	0,000000	0,000288	0,000000	0,000003	0,000000	0,000000
Ч	0,001850	0,000000	0,000006	0,000000	0,000000	0,005281	0,000003	0,000003	0,002022	0,000000	0,000113	0,000051	0,000000	0,001172	0,000006	0,000000	0,000068	0,000003	0,002819	0,000172	0,000000	0,000000	0,000000	0,000003	0,000208	0,000000	0,000000	0,000065	0,000000	0,000000	0,000000
Ш	0,000580	0,000003	0,000033	0,000000	0,000003	0,002025	0,000000	0,000000	0,001600	0,000000	0,000211	0,000511	0,000009	0,000306	0,000167	0,000003	0,000003	0,000009	0,000556	0,000181	0,000003	0,000003	0,000000	0,000000	0,000000	0,000000	0,000000	0,000277	0,000000	0,000000	0,000000
Щ	0,000419	0,000000	0,000000	0,000000	0,000000	0,002397	0,000000	0,000000	0,001344	0,000000	0,000000	0,000000	0,000000	0,000042	0,000000	0,000000	0,000024	0,000003	0,000000	0,000137	0,000000	0,000000	0,000000	0,000000	0,000000	0,000000	0,000000	0,000000	0,000000	0,000000	0,000000
Ы	0,000113	0,000488	0,001588	0,000181	0,000401	0,003152	0,000110	0,000274	0,000708	0,001195	0,000425	0,001704	0,001793	0,001629	0,000476	0,001011	0,000443	0,001433	0,000963	0,000190	0,000039	0,002750	0,000003	0,000303	0,000604	0,000003	0,000000	0,000000	0,000098	0,000000	0,000051
Ь	0,000137	0,000503	0,001112	0,000398	0,000467	0,000505	0,000134	0,000544	0,000723	0,000000	0,001088	0,000208	0,000470	0,000482	0,000824	0,000306	0,002536	0,000565	0,000184	0,000056	0,000048	0,000056	0,000336	0,001493	0,000018	0,000000	0,000000	0,000265	0,000464	0,000256	
Э	0,000003	0,000006	0,000024	0,000009	0,000003	0,000000	0,000000	0,000000	0,000000	0,000006	0,001374	0,000155	0,000027	0,000095	0,000000	0,000074	0,000045	0,000054	0,002530	0,000000	0,000149	0,000003	0,000003	0,000000	0,000006	0,000000	0,000000	0,000000	0,000000	0,000000	0,000000
Ю	0,000051	0,000220	0,000235	0,000071	0,000815	0,000039	0,000101	0,000190	0,000193	0,000036																					

Після обрахування частот моно- та біграм, обчислюємо ентропію на символ джерела та надлишковість.

Таблица 3. Отримані значення етропій

	Текст з пробілами	Текст без пробілів
Монограми	4.422432424271199	4.459897596950714
Біграми з перетинами	3.9870253553653483	4.107407950870605
Біграми без перетинів	3.986834015524383	4.107209868359202

За формулою $R = 1 - \frac{H_{\infty}}{H_0}$ обраховуємо надлишковість джерела ВТ

Таблица 4. Отримані значення надлишковості

	Текст з пробілами	Текст без пробілів
Монограми	0.11552	0.10802
Біграми з перетинами	0.2026	0.17852
Біграми без перетинів	0.20264	0.17856

Результати роботи CoolPinkProgram.

$H^{(10)}$

Лабораторная работа №1

Произвольная часть текста:
андарту поведения о котором по его мнению знает другой человек и тот другой

Использованные буквы:

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

Введенный символ: o

Символ по счету: 1

Номер эксперимента: 50

Поле ввода символов:
o

Продолжить Другой

Неравенство для энтропии:
1.45831222203456 < H < 2.20605318311958

Двоичная таблица угаданных символов:
10000000000000000000000000000000
10000000000000000000000000000000
10000000000000000000000000000000
00100000000000000000000000000000
00001000000000000000000000000000

Вероятности

q[1] = 0.56
q[2] = 0.18
q[3] = 0.04
q[4] = 0.04
q[5] = 0.02
q[6] = 0
q[7] = 0.02
q[8] = 0.06
q[9] = 0
q[10] = 0
q[11] = 0.02
q[12] = 0
q[13] = 0
q[14] = 0
q[15] = 0.02
q[16] = 0
q[17] = 0
q[18] = 0
q[19] = 0
q[20] = 0
q[21] = 0
q[22] = 0.02
q[23] = 0
q[24] = 0
q[25] = 0
q[26] = 0.02
q[27] = 0
q[28] = 0
q[29] = 0
q[30] = 0
q[31] = 0
q[32] = 0

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

Произвольная часть текста:		
нах_природы мы обычно подразумеваем такие вещи как силы тяготения или насле		
Использованные буквы:		

Порядок n-граммы:	Введенный символ: н	Неравенство для энтропии:
5 символов	Символ по счету: 1	1,67287299884748 < H < 2,19533621343214
10 символов		
15 символов	Номер эксперимента: 50	Двоичная таблица угаданных символов:
25 символов		10000000000000000000000000000000 ^
30 символов	Поле ввода символов:	10000000000000000000000000000000
35 символов	n	10000000000000000000000000000000
40 символов		00000000100000000000000000000000
45 символов	Продолжить	10000000000000000000000000000000 v
50 символов	Другой	

Строка состояния:

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

$$\mathbf{H}^{(30)}$$
[illegible]

Таблиця 5. Значення для CoolPinkProgram

1.45831	<	$H^{(10)}$	<	2.20605	0.55879	<	R	<	0.708338
1.67287	<	$H^{(20)}$	<	2.19534	0.560932	<	R	<	0.665426
1.49243	<	$H^{(30)}$	<	2.23341	0.553318	<	R	<	0.701514

Висновки: в даній роботі були вивчені та засвоєні поняття ентропії на символ джерела та надлишковості відкритого тексту. За допомогою засобів мови програмування ці значення були практично отримані. Значення вийшли приблизно однакові для ситуацій, коли пробіли вилучалися/не вилучалися з ВТ, а також коли біграми рахувалися із перетинами чи без. За отриманими даними надлишковість вийшла приблизно 20%, отже приблизно п'яту частину з усіх символів тексту російської мови можна вилучити, не втративши можливість зрозуміти текст.