

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

Комп'ютерний практикум №1
З дисципліни «Криптографія»

Виконали:
Студенти групи ФБ-33
Рудий А.О., Шкуропінський М.М.

Експериментальна оцінка ентропії на символ джерела відкритого тексту

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:

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.

Хід роботи:

1. Для виконання лабораторної роботи використали два тексти: «Тарас Бульба» та «Мертвые души» Миколи Гоголя.

Текст з пробілами:

Частота букв:

letter	frequency	ь	0,0171
	0,1628	я	0,0165
о	0,0940	ы	0,0162
е	0,0702	б	0,0159
а	0,0660	г	0,0151
и	0,0576	ч	0,0150
н	0,0517	з	0,0141
т	0,0503	ж	0,0091
с	0,0428	й	0,0088
в	0,0393	х	0,0082
л	0,0381	ш	0,0082
р	0,0352	ю	0,0053
к	0,0341	ц	0,0032
д	0,0260	щ	0,0024
у	0,0254	э	0,0018
м	0,0248	ф	0,0010
п	0,0233	ъ	0,0002
		ё	0,0001

Надлишквість: 0.13773755543420696

Ентропія: 4.3867281461341525

Частота біграм без перетенів Н₂:

	а	б	в	г	д	е	ё	ж	з	и	й	к	л	м	н	о
а	0,00001	0,00066	0,00286	0,00057	0,00141	0,00105	0,00001	0,00109	0,00324	0,00007	0,00064	0,00604	0,00651	0,00319	0,00394	0,00001
б	0,00127	0,00000	0,00004	0,00000	0,00003	0,00207	0,00000	0,00001	0,00000	0,00057	0,00000	0,00014	0,00079	0,00003	0,00026	0,00197
в	0,00554	0,00001	0,00006	0,00001	0,00025	0,00494	0,00000	0,00000	0,00036	0,00324	0,00000	0,00026	0,00055	0,00017	0,00092	0,00627
г	0,00086	0,00000	0,00001	0,00000	0,00109	0,00031	0,00000	0,00000	0,00000	0,00062	0,00000	0,00007	0,00118	0,00001	0,00028	0,00780
д	0,00447	0,00003	0,00082	0,00000	0,00001	0,00431	0,00000	0,00001	0,00001	0,00191	0,00000	0,00028	0,00052	0,00009	0,00179	0,00311
е	0,00003	0,00133	0,00175	0,00276	0,00255	0,00107	0,00000	0,00082	0,00105	0,00010	0,00184	0,00122	0,00564	0,00363	0,00570	0,00019
ё	0,00000	0,00000	0,00001	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000
ж	0,00090	0,00007	0,00000	0,00000	0,00062	0,00343	0,00000	0,00002	0,00000	0,00142	0,00000	0,00013	0,00001	0,00002	0,00088	0,00007
з	0,00487	0,00020	0,00085	0,00025	0,00090	0,00043	0,00000	0,00011	0,00002	0,00030	0,00000	0,00012	0,00017	0,00016	0,00156	0,00053
и	0,00007	0,00067	0,00248	0,00041	0,00134	0,00191	0,00000	0,00026	0,00193	0,00041	0,00114	0,00264	0,00412	0,00232	0,00293	0,00014
й	0,00000	0,00001	0,00000	0,00000	0,00015	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000	0,00019	0,00002	0,00006	0,00029	0,00000
к	0,00816	0,00000	0,00018	0,00000	0,00001	0,00060	0,00000	0,00004	0,00000	0,00339	0,00000	0,00000	0,00053	0,00001	0,00053	0,00949
л	0,00430	0,00005	0,00003	0,00014	0,00004	0,00337	0,00001	0,00026	0,00001	0,00589	0,00000	0,00040	0,00006	0,00001	0,00022	0,00580
м	0,00227	0,00004	0,00000	0,00003	0,00000	0,00342	0,00000	0,00000	0,00000	0,00298	0,00000	0,00007	0,00022	0,00001	0,00116	0,00263
н	0,00865	0,00001	0,00001	0,00002	0,00025	0,00087	0,00000	0,00005	0,00002	0,00095	0,00000	0,00041	0,00000	0,00000	0,00241	0,00883
о	0,00000	0,00348	0,00760	0,00437	0,00473	0,00197	0,00000	0,00180	0,00144	0,00074	0,00287	0,00190	0,00542	0,00464	0,00444	0,00015
п	0,00137	0,00000	0,00000	0,00000	0,00000	0,00207	0,00000	0,00000	0,00000	0,00098	0,00000	0,00018	0,00077	0,00000	0,00011	0,00883
р	0,00618	0,00009	0,00028	0,00014	0,00029	0,00522	0,00000	0,00020	0,00005	0,00450	0,00000	0,00033	0,00008	0,00018	0,00066	0,00706
с	0,00177	0,00005	0,00154	0,00002	0,00027	0,00357	0,00006	0,00001	0,00002	0,000129	0,00000	0,00352	0,00235	0,00071	0,00077	0,00228
т	0,00544	0,00002	0,00207	0,00001	0,00013	0,00496	0,00000	0,00000	0,00001	0,00291	0,00000	0,00048	0,00020	0,00005	0,00109	0,01221
у	0,00002	0,00073	0,00063	0,00114	0,00187	0,00020	0,00000	0,00161	0,00045	0,00001	0,00011	0,00096	0,00118	0,00113	0,00021	0,00000
ф	0,00021	0,00000	0,00000	0,00000	0,00000	0,00016	0,00000	0,00000	0,00000	0,00014	0,00000	0,00000	0,00002	0,00000	0,00000	0,00023
х	0,00074	0,00000	0,00025	0,00000	0,00000	0,00008	0,00000	0,00000	0,00000	0,00018	0,00000	0,00000	0,00026	0,00004	0,00014	0,00190
ц	0,00077	0,00000	0,00007	0,00000	0,00000	0,00075	0,00000	0,00000	0,00000	0,00013	0,00000	0,00005	0,00000	0,00000	0,00000	0,00018
ч	0,00177	0,00000	0,00000	0,00000	0,00000	0,00312	0,00000	0,00000	0,00000	0,00281	0,00000	0,00041	0,00003	0,00001	0,00064	0,00007
ш	0,00089	0,00000	0,00003	0,00000	0,00000	0,00186	0,00000	0,00000	0,00000	0,00207	0,00000	0,00065	0,00037	0,00002	0,00033	0,00026
щ	0,00033	0,00000	0,00000	0,00000	0,00000	0,00122	0,00000	0,00000	0,00000	0,00063	0,00000	0,00000	0,00000	0,00000	0,00003	0,00001
ъ	0,00000	0,00000	0,00000	0,00000	0,00000	0,00013	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000
ы	0,00000	0,00014	0,00083	0,00014	0,00008	0,00123	0,00000	0,00002	0,00003	0,00001	0,00150	0,00027	0,00211	0,00112	0,00024	0,00000
ь	0,00000	0,00013	0,00000	0,00006	0,00004	0,00061	0,00001	0,00000	0,00013	0,00005	0,00000	0,00132	0,00000	0,00027	0,00109	0,00001
э	0,00000	0,00000	0,00000	0,00000	0,00005	0,00000	0,00000	0,00000	0,00000	0,00000	0,00003	0,00007	0,00001	0,00001	0,00001	0,00000
ю	0,00000	0,00028	0,00000	0,00001	0,00024	0,00000	0,00000	0,00002	0,00000	0,00000	0,00000	0,00003	0,00002	0,00003	0,00004	0,00000
я	0,00000	0,00002	0,00033	0,00008	0,00010	0,00010	0,00000	0,00012	0,00025	0,00009	0,00016	0,00040	0,00063	0,00039	0,00074	0,00000
	0,00170	0,00696	0,01403	0,00401	0,00763	0,00323		0,00215	0,00433	0,01023	0,00001	0,00946	0,00226	0,00495	0,01498	0,00857

п	р	с	т	у	ф	х	ц	ч	ш	щ	ъ	ы	ь	э	ю	я
0,00093	0,00222	0,00303	0,00423	0,00009	0,00012	0,00094	0,00010	0,00081	0,00083	0,00023	0,00000	0,00000	0,00000	0,00000	0,00064	0,00179
0,00000	0,00142	0,00013	0,00002	0,00142	0,00000	0,00005	0,00001	0,00001	0,00002	0,00007	0,00007	0,00386	0,00005	0,00000	0,00000	0,00043
0,00020	0,00059	0,00278	0,00014	0,00062	0,00000	0,00003	0,00003	0,00005	0,00143	0,00001	0,00001	0,00244	0,00017	0,00000	0,00000	0,00027
0,00000	0,00066	0,00001	0,00001	0,00073	0,00000	0,00000	0,00000	0,00001	0,00001	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000
0,00011	0,00131	0,00042	0,00017	0,00181	0,00000	0,00005	0,00023	0,00002	0,00008	0,00000	0,00005	0,00048	0,00090	0,00000	0,00004	0,00031
0,00094	0,00513	0,00417	0,00466	0,00007	0,00000	0,00064	0,00063	0,00100	0,00085	0,00074	0,00000	0,00000	0,00000	0,00000	0,00026	0,00017
0,00000	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000
0,00000	0,00000	0,00003	0,00000	0,00019	0,00000	0,00000	0,00008	0,00002	0,00000	0,00000	0,00000	0,00000	0,00009	0,00000	0,00000	0,00000
0,00000	0,00014	0,00002	0,00002	0,00026	0,00000	0,00000	0,00001	0,00003	0,00001	0,00000	0,00004	0,00039	0,00015	0,00000	0,00000	0,00043
0,00020	0,00059	0,00233	0,00033	0,00001	0,00023	0,00165	0,00096	0,00227	0,00050	0,00018	0,00000	0,00000	0,00000	0,00000	0,00027	0,00106
0,00000	0,00000	0,00045	0,00026	0,00000	0,00000	0,00000	0,00002	0,00004	0,00013	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000
0,00000	0,00151	0,00007	0,00039	0,00221	0,00000	0,00000	0,00001	0,00000	0,00001	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000
0,00008	0,00000	0,00121	0,00008	0,00125	0,00001	0,00000	0,00000	0,00009	0,00000	0,00001	0,00000	0,00059	0,00372	0,00000	0,00082	0,00146
0,00004	0,00003	0,00011	0,00001	0,00200	0,00000	0,00000	0,00001	0,00002	0,00000	0,00001	0,00000	0,00077	0,00006	0,00000	0,00000	0,00043
0,00000	0,00004	0,00041	0,00032	0,00243	0,00001	0,00000	0,00025	0,00014	0,00000	0,00004	0,00000	0,00320	0,00139	0,00000	0,00013	0,00100
0,00122	0,00581	0,00533	0,00579	0,00008	0,00008	0,00050	0,00006	0,00172	0,00095	0,00016	0,00000	0,00000	0,00000	0,00002	0,00094	0,00053
0,00002	0,00611	0,00002	0,00006	0,00068	0,00000	0,00000	0,00004	0,00005	0,00003	0,00001	0,00000	0,00023	0,00007	0,00000	0,00000	0,00028
0,00010	0,00000	0,00027	0,00072	0,00258	0,00003	0,00018	0,00003	0,00008	0,00035	0,00003	0,00000	0,00130	0,00053	0,00000	0,00016	0,00101
0,00124	0,00016	0,00061	0,00965	0,00074	0,00001	0,00022	0,00002	0,00023	0,00008	0,00000	0,00003	0,00034	0,00255	0,00000	0,00020	0,00365
0,00011	0,00222	0,00097	0,00006	0,00140	0,00000	0,00001	0,00007	0,00012	0,00001	0,00000	0,00001	0,00146	0,00559	0,00000	0,00007	0,00045
0,00081	0,00063	0,00120	0,00147	0,00000	0,00001	0,00050	0,00001	0,00083	0,00084	0,00020	0,00000	0,00000	0,00000	0,00000	0,00097	0,00005
0,00000	0,00012	0,00001	0,00004	0,00002	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000	0,00001	0,00004	0,00000	0,00000	0,00000
0,00000	0,00017	0,00005	0,00003	0,00020	0,00000	0,00000	0,00000	0,00000	0,00001	0,00000	0,00001	0,00000	0,00000	0,00000	0,00000	0,00000
0,00000	0,00000	0,00000	0,00000	0,00017	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000	0,00027	0,00001	0,00000	0,00002	0,00001
0,00000	0,00006	0,00000	0,00039	0,00069	0,00000	0,00001	0,00000	0,00000	0,00013	0,00000	0,00000	0,00000	0,00025	0,00000	0,0	

	а	б	в	г	д	е	ё	ж	з	и	й	к	л	м	н	о	п
а	0,00001	0,00068	0,00287	0,00057	0,00143	0,00105	0,00000	0,00112	0,00330	0,00008	0,00070	0,00609	0,00653	0,00325	0,00401	0,00001	0,00092
б	0,00121	0,00000	0,00004	0,00001	0,00003	0,00205	0,00000	0,00001	0,00000	0,00059	0,00000	0,00014	0,00072	0,00003	0,00025	0,00198	0,00000
в	0,00554	0,00001	0,00007	0,00001	0,00026	0,00485	0,00000	0,00000	0,00038	0,00323	0,00000	0,00026	0,00052	0,00018	0,00099	0,00623	0,00019
г	0,00082	0,00000	0,00001	0,00000	0,00101	0,00033	0,00000	0,00000	0,00000	0,00059	0,00000	0,00008	0,00114	0,00002	0,00029	0,00782	0,00000
д	0,00446	0,00003	0,00085	0,00001	0,00001	0,00438	0,00000	0,00001	0,00002	0,00191	0,00000	0,00027	0,00053	0,00008	0,00175	0,00308	0,00010
е	0,00003	0,00135	0,00170	0,00279	0,00265	0,00109	0,00000	0,00076	0,00107	0,00011	0,00181	0,00118	0,00550	0,00365	0,00575	0,00020	0,00097
ё	0,00000	0,00000	0,00001	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000
ж	0,00089	0,00006	0,00000	0,00000	0,00066	0,00346	0,00000	0,00001	0,00000	0,00140	0,00000	0,00013	0,00001	0,00001	0,00087	0,00008	0,00000
з	0,00492	0,00017	0,00084	0,00025	0,00086	0,00042	0,00000	0,00012	0,00002	0,00029	0,00000	0,00011	0,00018	0,00017	0,00162	0,00054	0,00000
и	0,00007	0,00063	0,00255	0,00042	0,00136	0,00189	0,00000	0,00026	0,00194	0,00039	0,00114	0,00271	0,00412	0,00241	0,00295	0,00014	0,00020
й	0,00000	0,00001	0,00000	0,00000	0,00015	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000	0,00020	0,00002	0,00006	0,00030	0,00000	0,00000
к	0,00801	0,00000	0,00017	0,00000	0,00000	0,00059	0,00000	0,00004	0,00000	0,00336	0,00000	0,00001	0,00053	0,00000	0,00051	0,00950	0,00000
л	0,00434	0,00004	0,00004	0,00015	0,00004	0,00341	0,00001	0,00026	0,00001	0,00597	0,00000	0,00042	0,00006	0,00001	0,00024	0,00575	0,00007
м	0,00224	0,00003	0,00000	0,00002	0,00000	0,00345	0,00000	0,00000	0,00000	0,00295	0,00000	0,00008	0,00022	0,00002	0,00114	0,00269	0,00004
н	0,00862	0,00001	0,00001	0,00002	0,00021	0,00867	0,00000	0,00004	0,00002	0,00680	0,00000	0,00040	0,00000	0,00000	0,00246	0,00878	0,00000
о	0,00000	0,00347	0,00768	0,00435	0,00473	0,00198	0,00000	0,00180	0,00141	0,00077	0,00280	0,00192	0,00554	0,00460	0,00432	0,00014	0,00122
п	0,00134	0,00000	0,00000	0,00000	0,00000	0,00206	0,00000	0,00000	0,00000	0,00096	0,00000	0,00018	0,00074	0,00000	0,00010	0,00880	0,00002
р	0,00609	0,00010	0,00026	0,00015	0,00028	0,00522	0,00000	0,00021	0,00004	0,00449	0,00000	0,00034	0,00008	0,00019	0,00068	0,00695	0,00011
с	0,00171	0,00005	0,00150	0,00002	0,00027	0,00348	0,00005	0,00001	0,00002	0,00130	0,						

р	с	т	у	ф	х	ц	ч	ш	щ	ъ	ы	ь	э	ю	я	
0,00227	0,00305	0,00421	0,00008	0,00012	0,00093	0,00010	0,00079	0,00082	0,00022	0,00000	0,00000	0,00000	0,00000	0,00065	0,00182	0,01446
0,00140	0,00014	0,00001	0,00144	0,00000	0,00005	0,00001	0,00001	0,00002	0,00009	0,00008	0,00388	0,00005	0,00000	0,00000	0,00045	0,00032
0,00057	0,00277	0,00014	0,00060	0,00000	0,00002	0,00003	0,00006	0,00137	0,00001	0,00001	0,00241	0,00018	0,00000	0,00000	0,00028	0,00583
0,00068	0,00001	0,00001	0,00076	0,00000	0,00000	0,00000	0,00002	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000	0,00067
0,00129	0,00042	0,00018	0,00183	0,00000	0,00004	0,00022	0,00002	0,00007	0,00000	0,00004	0,00051	0,00086	0,00000	0,00003	0,00030	0,00121
0,00510	0,00427	0,00461	0,00008	0,00000	0,00067	0,00060	0,00099	0,00083	0,00076	0,00000	0,00000	0,00000	0,00000	0,00026	0,00015	0,01722
0,00000	0,00000	0,00001	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000	0,00006
0,00000	0,00003	0,00000	0,00020	0,00000	0,00000	0,00008	0,00002	0,00000	0,00000	0,00000	0,00000	0,00009	0,00000	0,00000	0,00000	0,00054
0,00013	0,00003	0,00002	0,00027	0,00000	0,00000	0,00001	0,00003	0,00001	0,00000	0,00004	0,00037	0,00016	0,00000	0,00000	0,00044	0,00132
0,00060	0,00232	0,00352	0,00001	0,00023	0,00166	0,00092	0,00228	0,00047	0,00021	0,00000	0,00000	0,00000	0,00000	0,00026	0,00109	0,01757
0,00000	0,00047	0,00026	0,00000	0,00000	0,00000	0,00002	0,00004	0,00012	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000	0,00659
0,00155	0,00009	0,00039	0,00224	0,00000	0,00000	0,00002	0,00000	0,00001	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000	0,00514
0,00000	0,00122	0,00008	0,00120	0,00001	0,00000	0,00000	0,00008	0,00000	0,00001	0,00000	0,00059	0,000370	0,00000	0,00083	0,00151	0,00588
0,00003	0,00011	0,00001	0,00197	0,00000	0,00000	0,00001	0,00002	0,00000	0,00001	0,00000	0,00075	0,00006	0,00000	0,00000	0,00040	0,00709
0,00005	0,00038	0,00030	0,00240	0,00001	0,00000	0,00023	0,00014	0,00000	0,00005	0,00000	0,00313	0,00138	0,00000	0,00014	0,00110	0,00338
0,00581	0,00539	0,00577	0,00008	0,00009	0,00048	0,00006	0,00170	0,00092	0,00017	0,00000	0,00000	0,00000	0,00002	0,00095	0,00052	0,01984
0,00621	0,00002	0,00006	0,00069	0,00000	0,00000	0,00005	0,00005	0,00003	0,00001	0,00000	0,00022	0,00007	0,00000	0,00000	0,00027	0,00010
0,00001	0,00027	0,00071	0,00261	0,00002	0,00018	0,00003	0,00007	0,00034	0,00003	0,00000	0,00134	0,00054	0,00000	0,00015	0,00101	0,00068
0,00016	0,00060	0,00964	0,00076	0,00001	0,00021	0,00002	0,00007	0,00000	0,00003	0,00003	0,00035	0,00254	0,00000	0,00017	0,00364	0,00259
0,00221	0,00098	0,00006	0,00140	0,00000	0,00001	0,00006	0,00013	0,00001	0,00000	0,00001	0,00142	0,00566	0,00000	0,00007	0,00046	0,00515
0,00062	0,00116	0,00145	0,00000	0,00001	0,00049	0,00001	0,00083	0,00082	0,00020	0,00000	0,00000	0,00000	0,00000	0,00097	0,00004	0,00621
0,00012	0,00001	0,00003	0,00002	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000	0,00001	0,00003	0,00000	0,00000	0,00000	0,00001
0,00017	0,00006	0,00003	0,00019	0,00000	0,00000	0,00000	0,00000	0,00001	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000	0,00362
0,00000	0,00000	0,00000	0,00019	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000	0,00028	0,00001	0,00000	0,00002	0,00000	0,00057
0,00006	0,00000	0,00339	0,00068	0,00000	0,00001	0,00000	0,00000	0,00013	0,00000	0,00000	0,00000	0,00023	0,00000	0,00000	0,00000	0,00066
0,00000	0,00000	0,00005	0,00027	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000	0,00050	0,00000	0,00000	0,00000	0,00016	
0,00001	0,00000	0,00000	0,00008	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000	0,00003	0,00000	0,00000	0,00000	0,00002
0,00000	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000	0,00010	0,00000
0,00030	0,00055	0,00054	0,00001	0,00000	0,00094	0,00002	0,00012	0,00043	0,00002	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000	0,00435
0,00000	0,00102	0,00013	0,00000	0,00000	0,00000	0,00010	0,00004	0,00034	0,00001	0,00000	0,00000	0,00000	0,00000	0,00030	0,00057	0,00980
0,00000	0,00000	0,00149	0,00000	0,00000	0,00004	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000	0,00002
0,00007	0,00010	0,00033	0,00000	0,00001	0,00005	0,00000	0,00010	0,00009	0,00024	0,00000	0,00000	0,00000	0,00000	0,00003	0,00000	0,00321
0,00013	0,00039	0,00144	0,00000	0,00000	0,00021	0,00003	0,00018	0,00001	0,00014	0,00000	0,00000	0,00000	0,00000	0,00007	0,00009	0,00907

Надлишковість: 0.2193585557415797

Ентропія: 3.9714843400047113

Текст без пробілів:

Частота букв:

letter	frequency
о	0,11225
е	0,083863
а	0,07878
и	0,068856
н	0,061797
т	0,060067
с	0,051117
в	0,04693
л	0,045547
р	0,042043
к	0,040772
д	0,031081
у	0,030343
м	0,029608
п	0,027842
ь	0,020383
я	0,019733
ы	0,019331
б	0,019047
г	0,018083
ч	0,017904
з	0,016891
ж	0,010836
й	0,010469
х	0,009821
ш	0,009751
ю	0,006278
ц	0,003847
щ	0,002926
э	0,00219
ф	0,001217
ъ	0,000286
ё	0,000111

Надлишковість: 0.12057763434583202

Ентропія: 4.474028607030048

Частота біграм без перетенів Н₂:

	а	б	в	г	д	е	ё	ж	з	и	й	к	л	м	н	о
а	0,0002	0,0016	0,0054	0,0012	0,0027	0,0017	0,0000	0,0016	0,0046	0,0014	0,0009	0,0091	0,0085	0,0047	0,0070	0,0010
б	0,0015	0,0000	0,0001	0,0000	0,0001	0,0026	0,0000	0,0000	0,0000	0,0008	0,0000	0,0002	0,0009	0,0000	0,0004	0,0026
в	0,0072	0,0003	0,0005	0,0004	0,0008	0,0062	0,0000	0,0001	0,0007	0,0045	0,0000	0,0009	0,0008	0,0004	0,0019	0,0085
г	0,0011	0,0000	0,0001	0,0000	0,0014	0,0004	0,0000	0,0000	0,0001	0,0008	0,0000	0,0001	0,0014	0,0000	0,0005	0,0101
д	0,0057	0,0001	0,0012	0,0001	0,0001	0,0056	0,0000	0,0000	0,0001	0,0026	0,0000	0,0005	0,0007	0,0001	0,0024	0,0040
е	0,0002	0,0028	0,0041	0,0041	0,0046	0,0018	0,0000	0,0012	0,0020	0,0014	0,0023	0,0025	0,0074	0,0056	0,0092	0,0015
ё	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
ж	0,0011	0,0001	0,0001	0,0000	0,0009	0,0045	0,0000	0,0000	0,0000	0,0018	0,0000	0,0002	0,0000	0,0001	0,0012	0,0002
з	0,0063	0,0003	0,0012	0,0004	0,0012	0,0005	0,0000	0,0002	0,0001	0,0004	0,0000	0,0003	0,0002	0,0003	0,0023	0,0008
и	0,0003	0,0018	0,0053	0,0012	0,0027	0,0030	0,0000	0,0005	0,0031	0,0018	0,0015	0,0048	0,0055	0,0037	0,0060	0,0016
й	0,0001	0,0004	0,0006	0,0003	0,0007	0,0001	0,0000	0,0002	0,0002	0,0006	0,0000	0,0009	0,0002	0,0004	0,0011	0,0003
к	0,0102	0,0005	0,0008	0,0002	0,0002	0,0008	0,0000	0,0002	0,0001	0,0047	0,0000	0,0003	0,0007	0,0002	0,0015	0,0123
л	0,0056	0,0004	0,0007	0,0004	0,0003	0,0047	0,0000	0,0004	0,0002	0,0082	0,0000	0,0010	0,0002	0,0002	0,0010	0,0082
м	0,0029	0,0003	0,0007	0,0003	0,0006	0,0045	0,0000	0,0001	0,0002	0,0046	0,0000	0,0006	0,0004	0,0003	0,0022	0,0040
н	0,0107	0,0003	0,0005	0,0001	0,0004	0,0109	0,0000	0,0001	0,0002	0,0090	0,0000	0,0007	0,0000	0,0001	0,0036	0,0114
о	0,0002	0,0057	0,0122	0,0060	0,0071	0,0031	0,0000	0,0030	0,0026	0,0023	0,0035	0,0036	0,0073	0,0066	0,0080	0,0016
п	0,0017	0,0000	0,0000	0,0000	0,0000	0,0027	0,0000	0,0000	0,0000	0,0012	0,0000	0,0002	0,0009	0,0000	0,0001	0,0112
р	0,0076	0,0001	0,0004	0,0002	0,0004	0,0067	0,0000	0,0002	0,0001	0,0058	0,0000	0,0005	0,0001	0,0003	0,0009	0,0089
с	0,0022	0,0002	0,0021	0,0001	0,0005	0,0045	0,0001	0,0001	0,0001	0,0018	0,0000	0,0047	0,0031	0,0010	0,0014	0,0031
т	0,0070	0,0003	0,0033	0,0002	0,0005	0,0066	0,0000	0,0002	0,0002	0,0040	0,0000	0,0010	0,0003	0,0003	0,0020	0,0158
у	0,0001	0,0012	0,0015	0,0016	0,0028	0,0004	0,0000	0,0021	0,0007	0,0007	0,0001	0,0018	0,0015	0,0017	0,0013	0,0003
ф	0,0002	0,0000	0,0000	0,0000	0,0000	0,0002	0,0000	0,0000	0,0000	0,0002	0,0000	0,0000	0,0000	0,0000	0,0000	0,0003
х	0,0009	0,0002	0,0007	0,0001	0,0003	0,0002	0,0000	0,0001	0,0001	0,0006	0,0000	0,0003	0,0005	0,0002	0,0006	0,0027
ц	0,0010	0,0000	0,0002	0,0000	0,0000	0,0010	0,0000	0,0000	0,0000	0,0002	0,0000	0,0001	0,0000	0,0000	0,0001	0,0003
ч	0,0024	0,0000	0,0001	0,0000	0,0000	0,0041	0,0000	0,0000	0,0000	0,0036	0,0000	0,0006	0,0000	0,0000	0,0010	0,0001
ш	0,0012	0,0000	0,0000	0,0000	0,0000	0,0024	0,0000	0,0000	0,0000	0,0027	0,0000	0,0009	0,0005	0,0000	0,0004	0,0004
щ	0,0004	0,0000	0,0000	0,0000	0,0000	0,0015	0,0000	0,0000	0,0000	0,0007	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
ъ	0,0000	0,0000	0,0000	0,0000	0,0000	0,0001	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
ы	0,0000	0,0005	0,0015	0,0003	0,0004	0,0017	0,0000	0,0001	0,0002	0,0005	0,0019	0,0007	0,0027	0,0015	0,0010	0,0003
ь	0,0002	0,0007	0,0013	0,0003	0,0006	0,0013	0,0000	0,0001	0,0005	0,0010	0,0000	0,0026	0,0002	0,0007	0,0026	0,0007
э	0,0000	0,0000	0,0000	0,0000	0,0001	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0001	0,0000	0,0000	0,0000	0,0000
ю	0,0001	0,0005	0,0003	0,0001	0,0005	0,0001	0,0000	0,0001	0,0001	0,0003	0,0000	0,0004	0,0001	0,0002	0,0004	0,0002
я	0,0001	0,0006	0,0016	0,0004	0,0012	0,0003	0,0000	0,0003	0,0006	0,0009	0,0002	0,0013	0,0009	0,0008	0,0023	0,0006

п	р	с	т	у	ф	х	ц	ч	ш	щ	ъ	ы	ь	э	ю	я
0,0029	0,0033	0,0056	0,0064	0,0006	0,0002	0,0013	0,0002	0,0017	0,0011	0,0003	0,0000	0,0000	0,0000	0,0003	0,0008	0,0025
0,0000	0,0018	0,0002	0,0000	0,0018	0,0000	0,0001	0,0000	0,0000	0,0000	0,0001	0,0001	0,0049	0,0001	0,0001	0,0000	0,0006
0,0009	0,0011	0,0044	0,0007	0,0009	0,0000	0,0001	0,0001	0,0003	0,0019	0,0000	0,0000	0,0030	0,0002	0,0002	0,0000	0,0004
0,0001	0,0008	0,0001	0,0000	0,0009	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
0,0003	0,0017	0,0007	0,0003	0,0024	0,0000	0,0001	0,0003	0,0001	0,0001	0,0000	0,0001	0,0006	0,0011	0,0000	0,0000	0,0004
0,0038	0,0069	0,0076	0,0070	0,0007	0,0001	0,0012	0,0009	0,0020	0,0011	0,0010	0,0000	0,0000	0,0000	0,0002	0,0003	0,0003
0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
0,0000	0,0000	0,0001	0,0001	0,0002	0,0000	0,0000	0,0001	0,0000	0,0000	0,0000	0,0000	0,0000	0,0001	0,0000	0,0000	0,0000
0,0002	0,0002	0,0002	0,0001	0,0004	0,0000	0,0000	0,0000	0,0001	0,0000	0,0000	0,0001	0,0005	0,0002	0,0000	0,0000	0,0006
0,0029	0,0013	0,0051	0,0056	0,0006	0,0004	0,0023	0,0012	0,0036	0,0008	0,0003	0,0000	0,0000	0,0000	0,0002	0,0003	0,0015
0,0008	0,0003	0,0015	0,0007	0,0002	0,0000	0,0001	0,0001	0,0004	0,0002	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0001
0,0006	0,0021	0,0008	0,0009	0,0030	0,0000	0,0001	0,0000	0,0003	0,0000	0,0000	0,0000	0,0000	0,0000	0,0001	0,0000	0,0001
0,0008	0,0001	0,0022	0,0004	0,0017	0,0000	0,0001	0,0000	0,0007	0,0000	0,0000	0,0000	0,0008	0,0046	0,0001	0,0011	0,0019
0,0010	0,0003	0,0010	0,0004	0,0027	0,0000	0,0001	0,0000	0,0005	0,0001	0,0000	0,0000	0,0010	0,0001	0,0001	0,0000	0,0005
0,0004	0,0002	0,0008	0,0006	0,0031	0,0000	0,0001	0,0003	0,0003	0,0000	0,0001	0,0000	0,0041	0,0018	0,0000	0,0002	0,0014
0,0039	0,0079	0,0092	0,0089	0,0008	0,0002	0,0009	0,0001	0,0032	0,0012	0,0003	0,0000	0,0000	0,0000	0,0003	0,0012	0,0009
0,0000	0,0080	0,0000	0,0001	0,0009	0,0000	0,0000	0,0001	0,0001	0,0000	0,0000	0,0000	0,0003	0,0001	0,0000	0,0000	0,0003
0,0002	0,0000	0,0004	0,0010	0,0032	0,0000	0,0002	0,0000	0,0001	0,0004	0,0000	0,0000	0,0017	0,0007	0,0000	0,0002	0,0013
0,0019	0,0003	0,0010	0,0124	0,0011	0,0000	0,0003	0,0000	0,0003	0,0001	0,0000	0,0000	0,0004	0,0032	0,0000	0,0002	0,0047
0,0007	0,0030	0,0018	0,0005	0,0020	0,0000	0,0001	0,0001	0,0005	0,0000	0,0000	0,0000	0,0019	0,0071	0,0001	0,0001	0,0006
0,0018	0,0009	0,0022	0,0024	0,0001	0,0000	0,0007	0,0000	0,0015	0,0010	0,0003	0,0000	0,0000	0,0000	0,0001	0,0012	0,0001
0,0000	0,0002	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
0,0005	0,0004	0,0005	0,0002	0,0003	0,0000	0,0000	0,0000	0,0002	0,0001	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
0,0001	0,0000	0,0001	0,0000	0,0003	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0003	0,0000	0,0000	0,0000	0,0000
0,0001	0,0001	0,0001	0,0043	0,0009	0,0000	0,0000	0,0000	0,0000	0,0002	0,0000	0,0000	0,0000	0,0003	0,0000	0,0000	0,0000
0,0001	0,0000	0,0000	0,0001	0,0003	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0006	0,0000	0,0000	0,0000
0,0000	0,0000	0,0000	0,0000	0,0001	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0001
0,0008	0,0005	0,0012	0,0010	0,0002	0,0000	0,0012	0,0000	0,0003	0,0006	0,0000	0,0000	0,0000	0,0000	0,0001	0,0000	0,0001
0,0012	0,0003	0,0024	0,0009	0,0003	0,0000	0,0001	0,0001	0,0006	0,0004	0,0000	0,0000	0,0000	0,0000	0,0001	0,0004	0,0009
0,0000	0,0000	0,0000	0,0019	0,0000	0,0000	0,0001	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
0,0005	0,0002	0,0005	0,0006	0,0001	0,0000	0,0001	0,0000	0,0003	0,0001	0,0003	0,0000	0,0000	0,0000	0,0000	0,0000	0,0001
0,0012	0,0004	0,0016	0,0025	0,0003	0,0000	0,0004	0,0001	0,0006	0,0001	0,0002	0,0000	0,0000	0,0000	0,0001	0,0001	0,0002

п	р	с	т	у	ф	х	ц	ч	ш	щ	ъ	ы	ь	э	ю	я
0,0030	0,0033	0,0056	0,0064	0,0006	0,0002	0,0014	0,0002	0,0017	0,0012	0,0003	0,0000	0,0000	0,0000	0,0003	0,0008	0,0025
0,0000	0,0018	0,0002	0,0000	0,0018	0,0000	0,0001	0,0000	0,0000	0,0000	0,0001	0,0001	0,0049	0,0001	0,0001	0,0000	0,0006
0,0009	0,0010	0,0044	0,0007	0,0009	0,0000	0,0001	0,0001	0,0002	0,0018	0,0000	0,0000	0,0031	0,0002	0,0002	0,0000	0,0004
0,0001	0,0009	0,0001	0,0000	0,0010	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
0,0002	0,0017	0,0007	0,0003	0,0024	0,0000	0,0001	0,0003	0,0001	0,0001	0,0000	0,0001	0,0006	0,0011	0,0000	0,0000	0,0004
0,0037	0,0070	0,0076	0,0070	0,0007	0,0001	0,0012	0,0008	0,0020	0,0011	0,0010	0,0000	0,0000	0,0000	0,0002	0,0003	0,0004
0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
0,0000	0,0000	0,0001	0,0001	0,0003	0,0000	0,0000	0,0001	0,0000	0,0000	0,0000	0,0000	0,0000	0,0001	0,0000	0,0000	0,0000
0,0002	0,0002	0,0002	0,0001	0,0004	0,0000	0,0000	0,0000	0,0001	0,0000	0,0000	0,0001	0,0005	0,0002	0,0000	0,0000	0,0006
0,0029	0,0013	0,0051	0,0056	0,0006	0,0004	0,0024	0,0012	0,0036	0,0008	0,0003	0,0000	0,0000	0,0000	0,0002	0,0003	0,0015
0,0008	0,0003	0,0015	0,0007	0,0002	0,0000	0,0001	0,0001	0,0004	0,0002	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0001
0,0006	0,0021	0,0008	0,0009	0,0030	0,0000	0,0001	0,0000	0,0003	0,0000	0,0000	0,0000	0,0000	0,0000	0,0001	0,0000	0,0001
0,0008	0,0001	0,0022	0,0004	0,0017	0,0000	0,0001	0,0000	0,0007	0,0000	0,0000	0,0000	0,0007	0,0047	0,0001	0,0010	0,0020
0,0010	0,0003	0,0010	0,0004	0,0027	0,0000	0,0001	0,0000	0,0004	0,0000	0,0000	0,0000	0,0010	0,0001	0,0001	0,0000	0,0006
0,0004	0,0002	0,0009	0,0006	0,0032	0,0000	0,0001	0,0003	0,0003	0,0000	0,0001	0,0000	0,0040	0,0018	0,0000	0,0002	0,0014
0,0040	0,0080	0,0091	0,0088	0,0008	0,0002	0,0009	0,0001	0,0032	0,0013	0,0002	0,0000	0,0000	0,0000	0,0003	0,0012	0,0009
0,0000	0,0079	0,0000	0,0001	0,0009	0,0000	0,0000	0,0001	0,0001	0,0000	0,0000	0,0000	0,0003	0,0001	0,0000	0,0000	0,0003
0,0002	0,0000	0,0004	0,0009	0,0033	0,0000	0,0002	0,0000	0,0001	0,0004	0,0000	0,0000	0,0017	0,0007	0,0000	0,0002	0,0013
0,0019	0,0003	0,0011	0,0125	0,0011	0,0000	0,0003	0,0000	0,0004	0,0001	0,0000	0,0000	0,0004	0,0032	0,0000	0,0002	0,0046
0,0007	0,0029	0,0018	0,0005	0,0019	0,0000	0,0001	0,0001	0,0005	0,0000	0,0000	0,0000	0,0018	0,0072	0,0001	0,0001	0,0007
0,0018	0,0009	0,0022	0,0024	0,0001	0,0000	0,0007	0,0000	0,0015	0,0011	0,0003	0,0000	0,0000	0,0000	0,0001	0,0012	0,0001
0,0000	0,0002	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
0,0005	0,0003	0,0006	0,0002	0,0004	0,0000	0,0000	0,0000	0,0001	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
0,0001	0,0000	0,0001	0,0000	0,0003	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0003	0,0000	0,0000	0,0000	0,0000
0,0001	0,0001	0,0001	0,0043	0,0009	0,0000	0,0000	0,0000	0,0000	0,0002	0,0000	0,0000	0,0000	0,0003	0,0000	0,0000	0,0000
0,0001	0,0000	0,0000	0,0001	0,0003	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0006	0,0000	0,0000	0,0000
0,0000	0,0000	0,0000	0,0000	0,0001	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0001
0,0008	0,0005	0,0012	0,0010	0,0002	0,0000	0,0013	0,0000	0,0003	0,0006	0,0000	0,0000	0,0000	0,0000	0,0001	0,0000	0,0001
0,0012	0,0003	0,0024	0,0009	0,0003	0,0000	0,0001	0,0001	0,0006	0,0005	0,0000	0,0000	0,0000	0,0000	0,0002	0,0004	0,0009
0,0000	0,0000	0,0000	0,0019	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
0,0004	0,0002	0,0005	0,0006	0,0001	0,0000	0,0001	0,0000	0,0003	0,0002	0,0003	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
0,0012	0,0004	0,0016	0,0025	0,0003	0,0000	0,0004	0,0001	0,0007	0,0001	0,0002	0,0000	0,0000	0,0000	0,0001	0,0001	0,0002

Ентропія: 4.145623071496935

Ентропія: 4.145623071496935

$$30,3\% < R^{(10)} < 42,7\%$$

[illegible]

$$43,7\% < R^{(20)} < 59,7\%$$

[illegible]

$$1,825 < H^{(30)} < 2,504$$

$$50,3\% < R^{(30)} < 63,8\%$$

Лабораторная работа №1

Произвольная часть текста:

руд_сравнить_учения_о_морали_

Использованные буквы:

Порядок n-граммы:

5 символов

10 символов

15 символов

20 символов

25 символов

30 символов

35 символов

40 символов

45 символов

50 символов

Введенный символ:

Символ по счету:

Номер эксперимента:

51

Неравенство для энтропии:

1,82597489005676 < H < 2,50444687389046

Двоичная таблица угаданных символов:

00010000000000000000000000000000

0000000000000000000000000100000000

10000000000000000000000000000000

10000000000000000000000000000000

01000000000000000000000000000000

Поле ввода символов:

Продолжить

Другой

Вероятности:

q[1] = 0,56

q[2] = 0,1

q[3] = 0,06

q[4] = 0,04

q[5] = 0

q[6] = 0

q[7] = 0

q[8] = 0

q[9] = 0

q[10] = 0

q[11] = 0,02

q[12] = 0,02

q[13] = 0,04

q[14] = 0

q[15] = 0

q[16] = 0

q[17] = 0,02

q[18] = 0,04

q[19] = 0

q[20] = 0,02

q[21] = 0

q[22] = 0

q[23] = 0,02

q[24] = 0,02

q[25] = 0,02

q[26] = 0

q[27] = 0

q[28] = 0,02

q[29] = 0

q[30] = 0

q[31] = 0

q[32] = 0

Строка состояния:

3. Надлишковість:

	Текст з пробілами	Текст без пробілів
H1	0.137737	0.120577
H2 (перетин)	0.219344	0.185171
H2 (без перетину)	0.219358	0.185129

Висновок:

У ході виконання лабораторної роботи, ми ознайомилися з поняттям ентропії, навчилися визначати її та також за допомогою її значень - надлишковість мови. Ми дізналися, що для тексту без пробілів значення виявилися меншими, також з 2 п. що зі зростанням довжини контексту ентропія зменшується.

Загалом, робота продемонструвала як на практиці вимірюється ентропія та надлишковість, і показала, що російська мова має помітну надлишковість.