

**МІНІСТЕРСТВО ОСВІТИ І НАУКИ, МОЛОДІ ТА СПОРТУ УКРАЇНИ
НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ УКРАЇНИ
«КИЇВСЬКИЙ ПОЛІТЕХНІЧНИЙ ІНСТИТУТ»**

ФІЗИКО- ТЕХНІЧНИЙ ІНСТИТУТ

Кафедра інформаційної безпеки

КОМП'ЮТЕРНИЙ ПРАКТИКУМ №1

з дисципліни

Криптографія

З теми: « Експериментальна оцінка ентропії на символ джерела відкритого тексту »

Виконав студент групи ФБ-91

Свищо Максим Іванович

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.

Завдання:

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 заміни 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:

Для тексту була взята одна з книжок про Гаррі Поттера і збережена вона у файлі text2.txt

```
Entropy with spaces: for letters - 4.402123312043531, for bigrams - 4.012884986309646, for bigrams with step 2 - 4.012673124538638
Entropy without spaces: for letters - 4.484090615985383, for bigrams - 4.1811174494425, for bigrams with step 2 - 4.179510576963886
```

Для тексту з пробілами:

Частоти літер:

$H_1 = 4.402123$

Надлишковість: 0,1347

1	Letter	Frequency	18	г	0,018794
2		0,159436	19	я	0,016434
3	о	0,092857	20	з	0,015564
4	а	0,069884	21	ь	0,015017
5	е	0,060188	22	ы	0,014766
6	и	0,059404	23	б	0,013728
7	н	0,054769	24	ч	0,012561
8	т	0,047779	25	ж	0,008366
9	с	0,045073	26	й	0,007946
10	р	0,045024	27	ш	0,007353
11	л	0,044638	28	х	0,006797
12	в	0,033669	29	ё	0,004341
13	м	0,027805	30	ю	0,004123
14	к	0,027785	31	щ	0,002555
15	д	0,02692	32	э	0,002336
16	у	0,025665	33	ф	0,00228
17	п	0,023835	34	ц	0,002099
18	г	0,018794	35	ъ	0,000205

Частоти біграм:

$H_2 = 4.01288$

Надлишковість: 0,21122

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	ъ	ы	ь	э	ю	я	з		
2.1E-05	0.00085	0.00271	0.00117	0.00232	0.00111	0.00089	0.00427	8.3E-05	0.00059	0.00394	0.00878	0.00368	0.00308	3.7E-05	0.00088	0.00448	0.00344	0.00389	8.5E-05	8.3E-05	0.00076	0.00014	0.00083	0.00059	0.0004	0	0	0	3E-06	0.00084	0.00169	3.6E-05	0.01802	
0.0006	0.00016	4.2E-05	3E-06	1E-05	0.00144	5.9E-06	4.5E-06	0.00078	0	0.00019	0.00115	5.3E-05	0.0006	0.00091	0	0.00148	0.00019	0.0013	1.5E-06	0.0011	0	0.3E-05	1.5E-06	0	7.4E-06	0.00015	0.00088	2.7E-05	7.4E-06	1.6E-05	0.00055	8.9E-06	0.00019	
0.00498	8.9E-06	4.3E-05	5.5E-05	0.00012	0.00141	0	0.00069	0.00079	0	0.00011	0.00053	0.00011	0.00099	0.00088	0.0001	0.00084	0.00017	0.00028	0.00054	0	2.8E-05	7.4E-06	4.3E-05	0.00059	0	1.5E-06	0.00033	0.00013	4.5E-06	0	0.00024	0.00019	0.00025	
0.00394	0	0.00014	0.0002	0.00104	0.00106	0	1.5E-06	0.0006	0	0.00013	0.00186	5.9E-06	0.00033	0.0007	3E-06	0.00018	6.1E-05	6.1E-05	0.00056	0	0	0	4.5E-05	0	0	0	0	0	0	0	2.2E-05	0	0.00066	
0.00455	4E-05	0.00089	1.2E-05	0.00013	0.00039	0.00153	1.9E-05	0.00203	0	0.00018	0.00082	7.6E-05	0.00163	0.00148	8.3E-05	0.00015	0.00022	0.0002	0.000162	0	4E-05	0.00014	1.5E-05	7E-05	0	7.4E-06	0.00053	0.00044	7.4E-06	3.6E-05	0.00067	0.00033	0.00159	
4E-05	0.00161	0.00118	0.00291	0.00297	0.00296	0.00069	0.00088	7.4E-05	0.00164	0.00086	0.0045	0.00376	0.00517	0.00014	0.00083	0.00705	0.00364	0.00445	9.5E-05	1.5E-05	0.00034	0.0002	0.00081	0.00086	0.00067	0	0	0	8.8E-05	0.00016	0.00061	0.01199		
0.00112	1.5E-05	0	1E-05	0.00068	0.00276	1.3E-05	0	0.00123	0	0.00013	2.7E-05	8.9E-06	0.00068	0.00044	0	3E-06	5.9E-06	3E-06	0.00024	0	0	1.3E-05	0	0	0	0	0	0	1.8E-05	0	0	0.00011	0.00085	
0.00629	0.00015	0.00071	0.00056	0.00086	0.00048	4.3E-05	0.00011	0.00035	0	0.00018	0.00701	0.00023	0.00159	0.00056	2.5E-05	0.00033	1.5E-06	4.5E-06	0.00037	0	0	1.5E-06	0	1E-05	0	1.6E-05	0.00029	0.00011	0	0	0.0002	8.9E-06	0.00138	
6.5E-05	0.00044	0.00198	0.00052	0.0024	0.00166	0.00027	0.00239	0.00056	0.00088	0.00199	0.00553	0.00269	0.00353	0.00106	0.00019	0.00122	0.00387	0.00037	0.00049	0.00018	0.00149	0.00077	0.00079	0.00049	0.00011	0	0	0	5.9E-06	0.00022	0.00105	7.4E-06	0.00132	
3E-06	0	0	1.5E-06	0.00012	1.5E-06	0	0	0	0	0.00011	3.9E-05	9.3E-05	0.00027	1.3E-05	3E-06	5.9E-06	0.00033	0.00028	0	3E-06	0	2.1E-05	0.00014	8.9E-05	0	0	0	0	0	0	0	0	0.00045	
0.00663	0.0001	0.00026	0.001	0	0.00048	5.9E-06	6.5E-05	0.00247	0	0.3E-05	0.0011	1E-05	0.00074	0.00737	1.5E-06	0.00188	0.00024	0.00064	0.00198	0	5E-05	5.2E-05	5.9E-06	1E-05	0	0	0	0	4E-05	0	0	0.00381		
0.00695	4E-05	1.5E-05	0.00014	0.00047	0.00313	0.00033	1.6E-05	0.00748	0	0.0004	0.00045	3E-06	0.00031	0.00351	2.7E-05	0	0.00166	4.2E-05	0.00144	0.00017	3E-06	0	0.00012	0.00042	0	0	0.00093	0.00368	4.6E-05	0.00077	0.00174	0.00026	0.00827	
0.00258	0.00087	1.5E-06	4.8E-05	1.1E-05	0.00029	0	0.00026	0.00013	5.2E-05	0.0013	0.00021	0.00031	0.00021	8.8E-05	0.00013	4.2E-05	0.00194	7.4E-06	1.5E-06	4.5E-06	4.5E-06	0	0.00026	7.6E-05	1.5E-06	0	0.00043	0.00011	0.00756	0.00011	0.00756			
0.01095	0	2.4E-05	0.0002	0.0007	0.00872	1.5E-05	1.5E-05	0.00764	0	0.0003	7.4E-06	1.5E-05	0.00018	0.00075	0	4E-05	0.00048	0.00071	0.00084	3.3E-05	0	0.00012	0.00018	1.5E-06	3E-05	1.5E-06	0.00098	0.00067	0	0.00014	0.00133	0.0002	0.00021	
1E-05	0.00011	0.00064	0.00084	0.0038	0.00147	0.00173	0.00019	0.00069	0.00315	0.0022	0.00622	0.00505	0.00892	0.00021	0.00149	0.00595	0.00551	0.00668	0.00026	0.00045	0.00053	7.3E-05	0.00023	0.00098	0.0002	0	4.9E-05	0.00015	0.00056	0.00015	0.01983			
0.00153	0	0.0036	0	0.00223	0	0	0	0.0107	0	0.0011	0.00065	0	0.00014	0.00095	0	0.00682	4.3E-05	9.8E-05	0.00077	0	0	3E-06	1.9E-05	5.9E-06	0	0	0.00043	2.5E-05	1.5E-05	0.0004	3E-05	6.1E-05		
0.00668	3.1E-05	0.00046	0.00031	0.00068	0.00125	0.00029	5.3E-05	0.00603	0	0.00027	7.7E-05	0.00115	0.0012	0.00883	8.2E-05	0.00287	0.00055	0.00075	0.00208	2.2E-05	0.00014	4.6E-05	0.00013	0.00019	2.4E-05	0	0.00133	0.00056	0.00012	0.00092	0.00018	0.00145		
0.00133	4.3E-05	0.001	3.9E-05	0.00025	0.00125	3.6E-05	1.8E-05	0.00152	0	0.00375	0.00284	0.00098	0.00112	0.00266	0.00204	0.00019	0.00129	0.00091	0.00067	2.4E-05	0.00016	3.7E-05	0.00032	7.4E-05	0	2.4E-05	0.00031	0.00306	3.6E-05	0.00016	0.00377	0.00052	0.00419	
0.00502	4.9E-05	0.00011	1.5E-06	7.7E-05	0.000478	0	0	0.00335	0	0.00066	0.00031	1.9E-05	0.00105	0.0119	0.0001	0.00078	0.0107	0.00022	0.00125	1.3E-05	4.5E-06	3.4E-05	0.00017	8.9E-06	1.2E-05	4.5E-06	0.00165	0.00052	1.5E-06	2.7E-05	0.00053	0.00032	0.00518	
3.1E-05	0.00047	0.00079	0.00077	0.00158	0.00018	0.00104	0.00023	0.00051	0.00014	0.00082	0.00244	0.00112	0.00054	2.1E-05	0.00056	0.00084	0.00195	0.00138	3E-06	4E-05	0.00048	4.5E-06	0.00104	0.00065	0.00024	0	0	7.4E-06	0.00094	7.3E-05	5.9E-06	0.00076		
0.00032	0	0	0	1.8E-05	0.00048	0	0	0.00035	0	0	7.9E-05	1.5E-06	1.5E-06	0.00023	0	0.00034	8.6E-06	3.3E-05	4.3E-05	0.00018	0	0	5.9E-06	1.2E-05	0	0	1.6E-05	0	0	0	0	0	4E-05	
0.00058	1.5E-06	0.00021	0	0	8E-05	0	0	0.00019	0	0	0.00022	0.00013	0.00009	0.00004	0	0.00014	4E-05	2.5E-05	0.00013	0	8.9E-06	0	0	1.2E-05	0	0	0	0	1.5E-06	0	0	0.00023		
0.00047	0	4.5E-05	0	1.5E-06	0.00055	0	0	0.0003	0	0	2.1E-05	0	0	0	0	0	0	1.2E-05	0.00013	0	0	0	0	0	0	0	0	0	0.00015	0	0	0	0.0002	
0.00166	4.5E-06	3E-06	0	0	0.00285	0.0001	0	0.00156	0	0.00018	7.7E-05	1.8E-06	0.00085	1E-04	0	0	0.0004	0.00036	5.9E-06	0	0	0.00032	0	0	0	0	0	0.00032	0	0	0.00014	9.7E-05		
0.00102	0	2.1E-05	0	0	0.00187	0	0	0.00136	0	0.00049	0.00068	2.1E-05	0.00026	0.00033	8.9E-06	5E-05	0	9.8E-05	0.00025	0	0	0	1.5E-06	3E-06	0	0	0	0	0.00056	1.5E-06	0	0.00028	3.6E-05	
0.00034	0	0	0	0	0.00068	0	0	0.00077	0	0	0	0	0	0	0	0	7.4E-05	0	0	0.00014	0	0	0	0	0	0	0	0	0	3.5E-05	0	0	0.00055	5.5E-06
0	0	0	0	0	4.9E-05	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1.5E-06	0.00011	1.5E-06	0	0	
0	0.00016	0.00079	0.00011	0.00011	0.00009	3.9E-05	8.3E-05	1.3E-05	0.00145	0.00015	0.0018	0.00115	0.00016	0	0.00016	0.00025	0.00048	0.00066	8.9E-06	0	0.00078	5.9E-06	0.00018	0.00049	5.9E-06	0	4.5E-06	0	0	1.2E-05	0	0	0.00465	
0	4.9E-05	1.1E-05	1.3E-05	3.1E-05	0.00024	0	0.00019	7.9E-05	0	0.00006	0	0.00021	0.00124	8.9E-06	7.4E-06	1.4E-05	0.00087	0.00021	0	9.5E-05	0	0.00011	4E-05	0.00061	1.2E-05	0	0	0	0.00031	0.00027	7E-05	0.00949		
0	3.3E-05	4.5E-06	4.5E-06	3E-06	0	0	0	0	1.6E-05	0.00014	0.00011	5.9E-06	6.1E-05	0	1.5E-06	6.8E-05	0.00018	0.00181	0	4.5E-06	1.5E-05	0	0	0	0	0	0	0	4.5E-06	0	0	0	5.5E-05	
7.4E-06	0.00014	2.4E-05	1.8E-05	0.00025	0	1.5E-05	1.2E-05	0	0	2.8E-05	6.9E-05	0.00014	4.5E-05	0	0.00014	4.8E-05	0.00013	0.00004	0	2.5E-05	2.1E-05	0.00015	1.2E-05	0.00041	0	0	0	0	1.8E-05	1.5E-06	1.5E-06	0.00021		
1.5E-06	2.5E-05	0.00046	0.00015	0.00097	0.00012	0.00011	0.00022	7.4E-06	3E-05	0.00016	0.00084	0.0004	0.00046	0	5.9E-05	0.00011	0.00057	0.00139	4.5E-06	0	0.00011	5.6E-05	0.00011	1.6E-05	0.00017	0	0	0	1.5E-06	0.00012	7.3E-05	0	0.00971	
0	1.2E-05	2.1E-05	6.1E-05	4.3E-05	0	1.8E-05	8.9E-05	0	2.1E-05	7.9E-05	0.00039	0.00045	0.00037	0	7.7E-05	0.00034	0.00017	0.00054	0	0	3.3E-05	0	1.5E-06	3E-05	0	0	0	0	0	0	0	0	0.00158	
0.00179	0.00032	0.01357	0.00753	0.00738	0.00332	0.00119	0.00487	0.00814	1.5E-06	0.00841	0.00217	0.00613	0.01564	0.01115	0.01673	0.00468	0.00193	0.00718	0.00526	0.00095	0.00177	0.00018	0.00512	0.00102	7E-05	0	0	0.00025	5.4E-05	0.00162	3E-06	0		

Частоти біграм з кроком 2:

N3=4.01267

Надлишковість: 0,21126

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	ъ
---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---

1	Letter	Frequency			
2	о	0,11047	18	я	0,019552
3	а	0,08314	19	з	0,018516
4	е	0,071605	20	ь	0,017866
5	и	0,070672	21	ы	0,017567
6	н	0,065157	22	б	0,016332
7	т	0,056842	23	ч	0,014943
8	с	0,053622	24	ж	0,009953
9	р	0,053564	25	й	0,009453
10	л	0,053105	26	ш	0,008748
11	в	0,040056	27	х	0,008086
12	м	0,03308	28	ё	0,005165
13	к	0,033055	29	ю	0,004905
14	д	0,032026	30	щ	0,003039
15	у	0,030533	31	э	0,002779
16	п	0,028356	32	ф	0,002712
17	г	0,022359	33	ц	0,002497
18	я	0,019552	34	ъ	0,000244

Частоти біграм:

Н2=4.18111

Надлишковість: 0,171138

	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	ъ	ь	э	ю	я	з		
а	0.00038	0.0016	0.00488	0.00272	0.00386	0.00187	0.00218	0.00558	0.00114	0.00071	0.00577	0.01074	0.00536	0.00584	0.00168	0.00053	0.00061	0.00623	0.00569	0.00007	0.00023	0.00108	0.00019	0.00156	0.00081	0.00048	0	0	0	0.00032	0.001	0.00227	4.4e-05	
б	0.00071	0.00019	5.5e-05	1.1e-05	1.6e-05	0.00171	7.1e-06	0.00096	0	0.00023	0.00137	8.1e-05	0.00072	0.00021	1.2e-05	0.00177	0.00016	1.8e-06	0.00133	0	4.1e-05	1.8e-06	0.00024	0.00018	0.00034	3.2e-05	8e-05	1.9e-05	0.00066	1.1e-05	0	0	0	
в	0.00599	0.00023	0.00039	0.00004	0.00063	0.00501	5.5e-05	0.00097	0.00387	0	0.00091	0.00075	0.00035	0.00162	0.00743	0.00097	0.00095	0.00025	0.0007	0.00079	2.1e-05	0.00015	3e-05	0.0002	0.00077	7.1e-06	1.8e-06	0.00016	1.8e-06	0.0001	0.00022	0	0	
г	0.00469	4.8e-05	0.00022	0.00026	0.00129	0.00127	5.5e-06	0.00076	0	0.0002	0.00022	2.3e-05	0.00046	0.00081	0.00166	0.00018	9.8e-05	0.00099	5.5e-06	3.5e-06	0	7.2e-05	3.5e-06	0	0	0	0	0	5.5e-06	2.7e-05	8.8e-06	0	0	
д	0.00544	8.1e-05	0.00119	0.00011	0.00034	0.00454	0.00183	9.5e-05	0.00236	0	0.00033	0.00102	0.00015	0.00119	0.00507	0.00026	0.0013	0.00044	0.00033	0.00021	8.8e-06	6.2e-05	0.00018	5.5e-05	8.7e-05	1.8e-06	8.8e-06	0.00063	0.00052	4.2e-05	4.2e-05	0.00081	0.00039	
е	0.00017	0.00263	0.00284	0.00099	0.00098	0.00092	0.00094	0.00176	0.00088	0.00195	0.00169	0.00558	0.00516	0.00752	0.0014	0.00287	0.00078	0.00068	0.00593	0.00029	8e-05	0.00065	0.00026	0.00144	0.00114	0.0008	0	0	0.00014	0.00011	0.00033	0.00073	0	
ж	0.00134	4.8e-05	8.7e-05	4.4e-05	0.00086	0.0003	1.8e-05	3.7e-05	0.00155	0	0.0002	4.1e-05	5.8e-05	0.00092	0.00098	0.00014	2.3e-05	0.00021	4.9e-05	0.00031	3.5e-06	1.1e-05	3.5e-06	3.9e-05	7.1e-06	0	0	2.1e-05	8.8e-06	0	1.2e-05	0.00014	0	
з	0.0075	0.00021	0.00102	0.00074	0.0011	0.0006	5.5e-05	0.00015	0.00045	0	0.00058	0.00086	0.00053	0.00056	0.00076	0.00022	0.00044	0.00017	8.5e-05	0.00047	5.5e-06	2.7e-05	3.5e-06	3.5e-05	3.5e-05	1.5e-05	0.00035	0.00013	2.7e-05	0	0.00025	1.1e-05	0	
и	0.0003	0.00134	0.00464	0.001	0.00039	0.00246	0.00042	0.00361	0.00192	0.00104	0.00346	0.00682	0.00403	0.0064	0.00281	0.00021	0.00115	0.00088	0.0049	0.00132	0.00031	0.00199	0.00089	0.00151	0.00077	0.00014	0	0	0.00021	0.00027	0.0015	1.1e-05	0	
й	7.1e-05	0.00022	0.00064	0.00045	0.00058	7.2e-05	8.5e-05	0.00023	0.00039	0	0.00065	0.00017	0.00048	0.00086	0.00039	0.00019	0.00039	0.00124	0.00061	0.00025	5.5e-05	9e-05	4.1e-05	0.00037	0.00021	8.8e-06	0	0	6.1e-05	0	7.1e-05	0	0	
к	0.00792	0.00021	0.00068	0.00033	0.00026	0.00065	6e-05	0.00016	0.00022	0	0.00024	0.00139	0.00018	0.00134	0.00909	0.00037	0.00113	0.00077	0.00099	0.00149	2.5e-05	0.0001	6.4e-05	0.00019	5.7e-05	1.8e-06	0	0	0.00013	1.8e-06	4.1e-05	0	0	
л	0.00835	0.00025	0.00072	0.00129	0.0017	0.00413	0.00043	0.00026	0.00931	0	0.00095	0.0007	0.00031	0.00127	0.00707	0.00072	0.00055	0.00188	0.00033	0.00189	0.00039	0.00011	1.1e-05	0.00061	0.00054	0	0	0.0011	0.00436	0.00015	0.00092	0.00211	0.00031	0
м	0.00365	0.00131	0.00075	0.00051	0.00041	0.00403	9.2e-05	0.00034	0.00549	1.8e-06	0.00084	0.00027	0.00037	0.00239	0.00445	0.00104	0.00035	0.00106	0.00039	0.00266	4.2e-05	7.6e-05	1.4e-05	0.0004	6.7e-05	7.1e-06	0	0.00114	9e-05	0.00011	8.8e-06	0.0006	0.00013	0
н	0.01306	0.00022	0.00045	0.0004	0.0009	0.01047	1.4e-05	0.00024	0.00943	0	0.00055	8.9e-05	0.00015	0.00047	0.01079	0.0007	0.0017	0.00112	0.00108	0.00045	6.2e-05	7.1e-05	0.00017	0.00034	2.1e-05	3.7e-05	1.8e-06	0.00394	0.00118	4.9e-05	0.00016	0.00146	0.00024	0
о	0.0002	0.00471	0.00929	0.00535	0.00547	0.00218	0.00131	0.00223	0.0172	0.00375	0.00364	0.00076	0.00696	0.01292	0.00247	0.00426	0.00918	0.00104	0.00062	0.00086	0.00012	0.00351	0.00019	0.00026	0	0	0	0	0.00043	0.00019	0.00096	0.00017	0	
п	0.00182	0	3.5e-06	3.5e-06	0	0.00265	0	0	0.0127	0	0.00014	0.00077	1.2e-05	0.00117	0.01131	8.1e-05	0.00788	6.5e-05	0.00012	0.00092	0	0	3.5e-06	2.5e-05	7.1e-06	0	0	0.00052	9e-05	3.4e-05	1.2e-05	0.00048	3.5e-05	0
р	0.0081	7.8e-05	0.00063	0.00046	0.00087	0.00617	0.00035	9.9e-05	0.00962	0	0.00038	0.00012	0.00132	0.00136	0.00058	0.00033	0.00357	0.00081	0.00097	0.00188	4.4e-05	0.00019	5.7e-05	0.0002	0.00023	2.8e-05	0	0.00158	0.00096	5.3e-05	0.00027	0.00112	0.00022	0
с	0.00166	0.0002	0.00148	0.00018	0.00061	0.00305	8e-05	0.00013	0.00137	0	0.00479	0.00347	0.00137	0.00193	0.00345	0.00287	0.00337	0.00181	0.01113	0.00123	0.00017	0.00024	3.9e-05	0.00048	0.00011	1.8e-06	2.8e-05	0.00037	0.00085	0.00011	0.00019	0.00432	0.00061	0
т	0.00606	0.00029	0.00036	0.00023	0.00038	0.00582	8.8e-05	0.00017	0.0043	0	0.0011	0.00046	0.00022	0.00187	0.01453	0.00048	0.00348	0.00184	0.00063	0.00165	5.5e-05	7.1e-05	4.6e-05	0.00047	3.2e-05	1.4e-05	5.5e-06	0.00197	0.00609	0.00012	3.4e-05	0.0007	0.00038	0
у	0.00017	0.00075	0.00158	0.00133	0.00025	0.00033	0.0013	0.00043	0.00112	0.00017	0.00139	0.00032	0.00161	0.00176	0.00055	0.0014	0.00117	0.00025	0.00202	0.00021	9.4e-05	0.00094	7.1e-06	0.0016	0.00081	0.00029	0	0	7.8e-05	0.00011	0.00017	7.1e-06	0	
ф	0.00038	0	3.5e-06	3.5e-06	2.1e-05	0.00038	0	0.00042	0	3.5e-06	9.4e-05	3.5e-06	8.8e-06	0.00039	7.1e-06	0.00041	1.8e-05	4.1e-05	3.1e-05	0.00022	0	7.1e-06	1.8e-05	0	1.8e-05	0	1.8e-05	0	1.8e-05	0	1.8e-06	0	0	
х	0.00072	9.7e-05	0.00053	0.00011	0.00015	0.00013	4.8e-05	0.00014	0.00043	0	0.00022	0.00033	0.00029	0.00057	0.00261	0.00044	0.00024	0.00039	0.00013	0.00026	3.4e-05	2.8e-05	3.5e-06	7.2e-05	5.3e-06	0	0	0	3.2e-05	1.8e-06	2.5e-05	0	0	
ц	0.00056	8.8e-06	6.5e-05	1.4e-05	1.1e-05	0.00066	0	7.1e-06	0.00037	0	4.6e-05	0	1.8e-05	0.00028	2.7e-05	8.8e-06	3.4e-05	0.00017	3.5e-06	1.8e-06	1.8e-06	0.00016	0	0	0.00017	0	0	0.00025	0	0	3.5e-06	0	0	
ч	0.00198	5.5e-06	8.8e-06	0	7.1e-06	0.00039	0.00012	5.5e-06	0.00186	0	0.0007	9.2e-05	1.8e-06	0.00102	0.00013	1.2e-05	4.8e-05	1.1e-05	0.00045	0.00067	7.1e-06	0	0	3.5e-06	0.00016	0	0	0.00087	0	1.8e-06	0.00041	0	0.00034	0
ш	0.00122	0	1.8e-05	3.5e-06	0	0.00033	0	0.00162	0	0.00059	0.00081	1.5e-05	0.00031	0.00039	1.8e-05	6e-05	3.5e-06	0.00012	0.00031	0	0	1.8e-06	3.5e-06	0	0	0	0	0.00087	0	1.8e-06	0	0.00034	0	0
щ	0.00004	0	0	0	0	0.00081	0	0	0.00092	0	1.8e-06	0	0	0	0	0	8.8e-06	1.8e-06	0	0.00017	0	0	0	0	0	0	0	0	0	3.9e-05	0	0	0.00066	0
ъ	0	0	0	0	5.5e-05	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.00018	0.00016	5.5e-06	0
ь	3.7e-05	0.00045	0.00146	0.00032	0.00041	0.00117	0.0001	0.00005	0.00034	0.00173	0.00041	0.00219	0.00155	0.00065	0.00078	0.00042	0.00108	0.00101	0.00013	1.4e-05	0.00102	1.4e-05	0.00038	0.0004	8.8e-06	5.3e-06	0	0	7.1e-05	1.8e-06	7.4e-05	0	0	
э	0.00012	0.00037	0.00111	0.00044	0.00053	0.00061	5.1e-05	0.00056	0.00068	0	0.00186	0.00017	0.00071	0.00256	0.00084	0.00103	0.00028	0.00028	0.00063	0.00034	0.00016	0.00018	0.00014	0.00061	0.00077	1.4e-05	0	0	0.00016	0.00037	0.00043	8.3e-05	0	
ю	0	3.9e-05	8.8e-06	7.1e-06	1.1e-05	0	0	8.8e-06	1.8e-06	1.8e-05	0.00016	0.00013	8.8e-06	8.1e-05	1.8e-06	8.8e-06	8.1e-05	0.00025	0	5.5e-06	2.5e-05	0	3.5e-06	0	0	0	0	0	0	3.5e-06	0	0	0	0
я	4.8e-05	0.00023	0.0002	0.00013	0.00016	0.000	0	8.8e-06	1.8e-05	1.8e-05	3.4e-05	0.00013	0.00015	0.00025	0.00012	0.00015	0.00025	0.00021	0.00025	0.00021	4.9e-05	0.00027	7.4e-05	0.00043	0.0001	0.0002	0	0	0.00015	0.00014	0.00018	0.00018	0	
з	3.5e-06	0.00016	0.0002	0.00013	0.0001	6.5e-05	5.5e-05	0.00017	7.4e-05	2.5e-05	0.00016	0.00012	0.00017	0.00005	0.00012	0.00027	0.00015	0.00036	0.00027	0.00027	5.1e-05	0	5.7e-05	1.8e-06	7.4e-05	6e-05	0	0	0	3.4e-05	0	1.6e-05	0	

Работа з CoolPinkProgram:

Лабораторная работа №1



Произвольная часть текста:
_подыскиваем_пояснения_только_нашему_плохому_поведению_только_наше_плохое_п

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: а
Символ по счету: 1
Номер эксперимента: 51
Поле ввода символов:
а
Продолжить Другой

Неравенство для энтропии:
3,3172425891973< H < 3,79140260879704
Двоичная таблица угаданных символов:
00010000000000000000000000000000
00001000000000000000000000000000
00000000000000000000000000000001
00000000100000000000000000000000
00000000000000000000000000000001
.....

Вероятности:
q[1] = 0,2745098
q[2] = 0,0588235
q[3] = 0,0980392
q[4] = 0,0588235
q[5] = 0,0196078
q[6] = 0,0588235
q[7] = 0
q[8] = 0
q[9] = 0,0392156
q[10] = 0,019607
q[11] = 0,019607
q[12] = 0
q[13] = 0
q[14] = 0
q[15] = 0,039215
q[16] = 0,039215
q[17] = 0,019607
q[18] = 0,039215
q[19] = 0,019607
q[20] = 0
q[21] = 0
q[22] = 0
q[23] = 0,019607
q[24] = 0
q[25] = 0,019607
q[26] = 0,039215
q[27] = 0,019607
q[28] = 0,039215
q[29] = 0
q[30] = 0
q[31] = 0
q[32] = 0,058823

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

Надлишковість: 0,2417186 < R10 < 0,336551

Лабораторная работа №1



Произвольная часть текста:
_что_случилось_нечто_непредвиденное_освобождающее_его_от_необходимости_выпо

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: о
Символ по счету: 1
Номер эксперимента: 51
Поле ввода символов:
о
Продолжить Другой

Неравенство для энтропии:
1,74402577714493< H < 2,44906801132392
Двоичная таблица угаданных символов:
10000000000000000000000000000000
00000100000000000000000000000000
10000000000000000000000000000000
00000000001000000000000000000000
10000000000000000000000000000000
.....

Вероятности:
q[1] = 0,5686274
q[2] = 0,1176470
q[3] = 0,0392156
q[4] = 0
q[5] = 0
q[6] = 0,0392156
q[7] = 0,0196078
q[8] = 0,0196078
q[9] = 0,0196078
q[10] = 0
q[11] = 0
q[12] = 0,039215
q[13] = 0
q[14] = 0
q[15] = 0
q[16] = 0,019607
q[17] = 0,019607
q[18] = 0
q[19] = 0,039215
q[20] = 0
q[21] = 0
q[22] = 0
q[23] = 0
q[24] = 0,019607
q[25] = 0
q[26] = 0
q[27] = 0
q[28] = 0,019607
q[29] = 0,019607
q[30] = 0
q[31] = 0
q[32] = 0

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

0,5101864 < R20 < 0,65119486

Произвольная часть текста:
 ловек_постоянно_каждую_секунду_находится_под_действием_нескольких_различных

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: y

Символ по счету: 1

Номер эксперимента: 51

Поле ввода символов:

Продолжить
Другой

Неравенство для энтропии:
 $2,12064035045134 < H < 2,7333716642707$

Двоичная таблица угаданных символов:

00000000000000000000000000000000	▲
10000000000000000000000000000000	
10000000000000000000000000000000	
00000010000000000000000000000000	
10000000000000000000000000000000	▼

Строка состояния:

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

$$0,453325668 < R_{30} < 0,57587194$$

Висновки: У ході виконання цієї лабораторної роботи я виміряв частоти повторювання символів і біграм у тексті і навчився вимірювати по цим даним ентропію і надлишковість мови. Було помічено, що у тексті без пробілів ентропія більше. Також доволі корисною і цікавою була робота з наданою програмою CoolPinkProgram.