

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

**КРИПТОГРАФІЯ**  
**КОМП’ЮТЕРНИЙ ПРАКТИКУМ №1**

Експериментальна оцінка ентропії на символ джерела  
відкритого текст

Виконав

Студенти гр.Фб-22 Пунько Артем

гр.Фб-22 Руденко Поліна

Київ – 2024

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

### **Постановка задачі:**

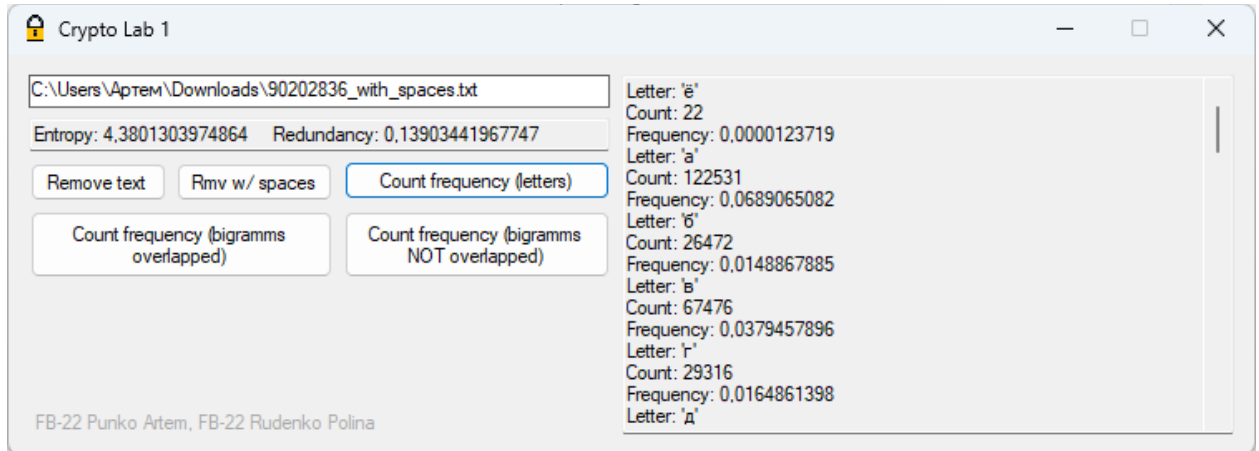
1. Написати програми для підрахунку частот букв і частот біграм в тексті, а також підрахунку  $H_1$  та  $H_2$  за безпосереднім означенням. Підрахувати частоти букв та біграм, а також значення  $H_1$  та  $H_2$  на довільно обраному тексті російською мовою достатньої довжини (щонайменше 1Мб), де імовірності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

### **Хід роботи:**

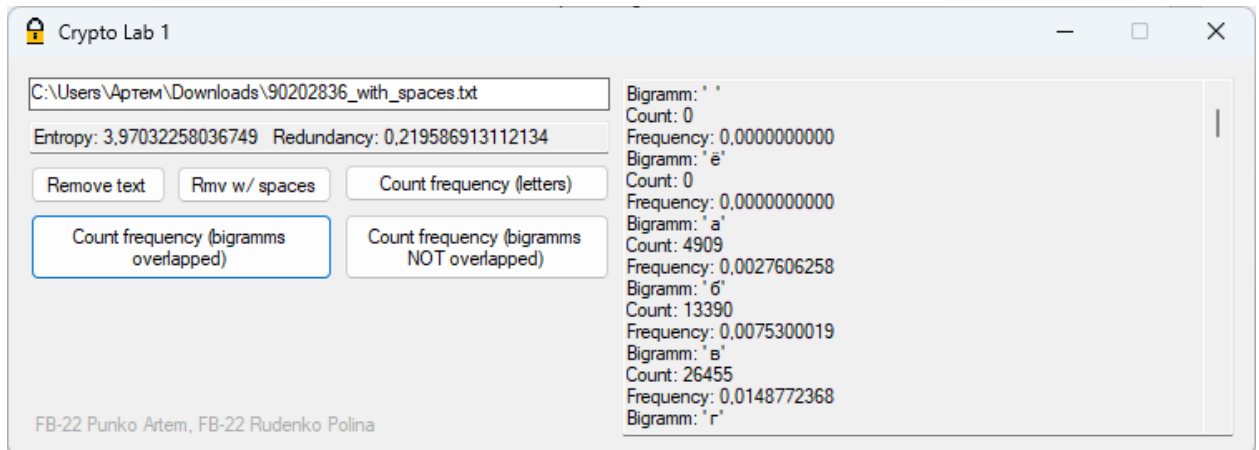
Щоб виконати лабораторну роботу було розроблено програму, для аналізу тексту та обчислення різних статистичних показників. Загальні етапи для роботи коду були: Зчитавання тексту з файлу, попередня обробка тексту, створення нової версії тексту без пробілів, підрахунок частот символів та біграм, обчислення ентропії  $H_1$  та  $H_2$ , обчислення надлишковості, збереження результатів у файлі.

# Аналіз тексту з пробілами

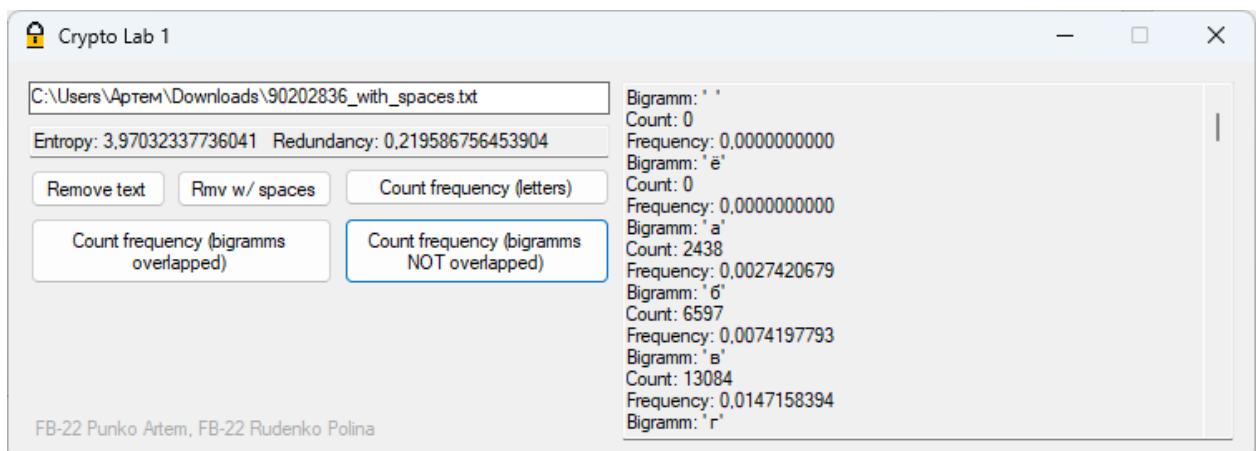
## Літери



## Біграми (з перекриттям)



## Біграми (без перекриття)



# Аналіз тексту БЕЗ пробілів

## Літери

Crypto Lab 1

C:\Users\Арте́м\Downloads\90202836\_without\_spaces.txt

Entropy: 4,46679772386303 Redundancy: 0,114502630410821

Remove text Rmv w/ spaces **Count frequency (letters)**

Count frequency (bigramms overlapped) Count frequency (bigramms NOT overlapped)

FB-22 Punko Artem, FB-22 Rudenko Polina

Letter: 'e'  
Count: 22  
Frequency: 0,0000147814  
Letter: 'a'  
Count: 122531  
Frequency: 0,0823264058  
Letter: 'b'  
Count: 26472  
Frequency: 0,0177860673  
Letter: 'b'  
Count: 67476  
Frequency: 0,0453359277  
Letter: 't'  
Count: 29316  
Frequency: 0,0196969005  
Letter: 'd'

## Біграми (з перекриттям)

Crypto Lab 1

C:\Users\Арте́м\Downloads\90202836\_without\_spaces.txt

Entropy: 4,14738686704838 Redundancy: 0,17782259496095

Remove text Rmv w/ spaces Count frequency (letters)

Count frequency (bigramms overlapped) Count frequency (bigramms NOT overlapped)

FB-22 Punko Artem, FB-22 Rudenko Polina

Bigramm: ' '  
Count: 0  
Frequency: 0,0000000000  
Bigramm: 'e'  
Count: 0  
Frequency: 0,0000000000  
Bigramm: 'a'  
Count: 0  
Frequency: 0,0000000000  
Bigramm: 'b'  
Count: 0  
Frequency: 0,0000000000  
Bigramm: 'b'  
Count: 0  
Frequency: 0,0000000000  
Bigramm: 'a'  
Count: 0  
Frequency: 0,0000000000  
Bigramm: 'r'

## Біграми (без перекриття)

Crypto Lab 1

C:\Users\Арте́м\Downloads\90202836\_without\_spaces.txt

Entropy: 4,14671433121088 Redundancy: 0,177955918373353

Remove text Rmv w/ spaces Count frequency (letters)

Count frequency (bigramms overlapped) **Count frequency (bigramms NOT overlapped)**

FB-22 Punko Artem, FB-22 Rudenko Polina

Bigramm: ' '  
Count: 0  
Frequency: 0,0000000000  
Bigramm: 'e'  
Count: 0  
Frequency: 0,0000000000  
Bigramm: 'a'  
Count: 0  
Frequency: 0,0000000000  
Bigramm: 'b'  
Count: 0  
Frequency: 0,0000000000  
Bigramm: 'a'  
Count: 0  
Frequency: 0,0000000000  
Bigramm: 'r'

## Таблиця з частотами літер з пробілом

а	122531	0,068907	
б	26472	0,014887	
в	67476	0,037946	
г	29316	0,016486	
д	45228	0,025434	
е	122639	0,068967	
ж	15493	0,008713	
з	25297	0,014226	
и	98672	0,055489	
й	17246	0,009698	
к	51478	0,028949	
л	75503	0,04246	
м	44468	0,025007	
н	97793	0,054995	
о	170265	0,09575	
п	38644	0,021732	
р	66198	0,037227	
с	78284	0,044024	
т	86506	0,048647	
у	41400	0,023282	
ф	2932	0,001649	
х	12873	0,007239	
ц	5323	0,002993	
ч	21686	0,012195	
ш	13757	0,007736	
щ	4236	0,002382	
ъ	659	0,000371	
ы	28753	0,01617	
ь	29561	0,016624	
э	4759	0,002676	
ю	9487	0,005335	
я	33399	0,018782	
	289865	0,163008	
Entropy	4,38013	Redundanc	0,139034

## Таблиця з частотами літер без пробілу

а	122531	0,082326		
б	26472	0,017786		
в	67476	0,045336		
г	29316	0,019697		
д	45228	0,030388		
е	122639	0,082399		
ж	15493	0,010409		
з	25297	0,016997		
и	98672	0,066296		
й	17246	0,011587		
к	51478	0,034587		
л	75503	0,050729		
м	44468	0,029877		
н	97793	0,065705		
о	170265	0,114398		
п	38644	0,025964		
р	66198	0,044477		
с	78284	0,052598		
т	86506	0,058122		
у	41400	0,027816		
ф	2932	0,00197		
х	12873	0,008649		
ц	5323	0,003576		
ч	21686	0,01457		
ш	13757	0,009243		
щ	4236	0,002846		
ъ	659	0,000443		
ы	28753	0,019319		
ь	29561	0,019862		
э	4759	0,003197		
ю	9487	0,006374		
я	33399	0,02244		
Entropy	4,466798	Redundanc	0,114503	

## Таблиця з частотами біграм з перекриттям 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	
а	0	0,002761	0,00753	0,014877	0,004695	0,007276	0,004204	0,001969	0,003968	0,011336	9E-06	0,010269	0,002603	0,005649	0,015527	0,01208	0,01525	0,003939	0,015572	0,006877	0,003882	0,000695	0,001186	0,000323	0,005507	0,0008
б	7,87E-06	0	0	0,12E-06	0,006429	0	0	0	0	0	0	0,562E-07	0,562E-07	0	0	0,112E-06	0	0	0,662E-07	0	0	0	0	0	0	0
в	0,016248	0,844E-06	0,000635	0,003148	0,000775	0,002012	0,000834	0,001086	0,003647	0,000125	0,000758	0,004373	0,008902	0,002668	0,004736	7,31E-06	0,000836	0,002778	0,00351	0,004808	0,00011	0,000462	0,000993	8,15E-05	0,000724	0,0012
г	0,000246	0	0,001045	1,69E-06	6,92E-05	1,01E-05	2,08E-05	0,002045	3,94E-06	0,000731	0	0,000227	0,000774	3,82E-05	0,000281	0,002158	0	0,001104	8,94E-05	6,19E-06	0,001148	0	6,35E-05	6,75E-06	2,02E-05	4,05E-06
д	0,006048	5,62E-07	0,00543	1,12E-05	3,43E-05	3,09E-05	0,000247	0,004276	5,62E-07	0,000463	0,002997	0	0,000128	0,00093	0,000118	0,001147	0,007129	0,000284	0,000571	0,003187	0,000196	0,000736	5,62E-07	4,22E-05	1,07E-05	4,16E-05
е	0,000751	0	0,000891	0	4,89E-05	0	0,000965	0,00051	1,12E-06	0	0,00069	0	8,44E-05	0,001458	3,94E-06	0,000235	0,000831	0	0,001159	6,19E-05	4,5E-06	0,00066	0	0	2,76E-05	4,5E-06
ж	0,001054	0	0,004121	3,61E-05	0,000887	1,69E-06	3,82E-05	0,003495	7,31E-06	3,37E-06	0,002425	5,62E-07	0,000214	0,000795	7,99E-05	0,001559	0,003568	8,04E-05	0,002112	0,003005	0,000278	0,001609	3,94E-06	5,01E-05	0,000148	2,76E-05
з	0,016948	0	2,98E-05	0,001079	0,001323	0,003465	0,00243	0,001585	0,000981	0,0012	0,000139	0,002443	0,000982	0,000698	0,003956	0,007302	0,000403	0,000896	0,000506	0,003656	0,003629	5,96E-05	1,12E-05	0,001009	0,000281	0,000907
и	0,000215	0	0,001281	6,86E-05	0	1,52E-05	0,000669	0,003643	0,002915	0	0,001219	0	8,66E-05	5,08E-06	8,44E-06	0,001125	5,57E-05	0	2,25E-06	1,74E-05	0	0,000164	0	0	5,62E-07	5,62E-05
й	0,001144	0	0,005026	0,000146	0,000713	0,000397	0,000689	0,000292	0,00137	6,41E-05	0,00316	0	0,000132	0,00098	0,000247	0,001606	0,000534	2,81E-05	0,000184	0,000136	1,57E-05	0,000254	0	0,50E-06	1,47E-05	8,44E-06
к	0,018751	0	8,38E-05	0,000394	0,002385	0,000474	0,001558	0,002322	0,000314	0,001731	0,000673	0,001289	0,001829	0,004731	0,002797	0,003005	0,000177	0,000688	0,000578	0,002775	0,003191	2,25E-06	2,76E-05	0,00154	0,0011	0,001185
л	0,007852	0	5,62E-07	1,07E-05	0	0	0,000153	5,06E-06	0	0	0	0	0,67E-05	9,62E-05	3,89E-05	0,000389	8,44E-06	1,12E-06	1,69E-05	0,000608	0,00027	0	2,81E-06	0	3,49E-05	5,51E-05
м	0,00048	5,62E-07	0,006975	0	0,000258	0	7,31E-06	0,000452	1,52E-05	1,12E-05	0,00237	0	2,92E-05	8,85E-05	3,94E-06	0,001814	0,000828	1,12E-06	0,001574	0,000107	0,000431	0,001426	0	5,62E-07	7,31E-06	1,69E-05
н	0,00985	5,62E-07	0,006735	2,25E-05	3,37E-06	0,000233	0,003746	0,000343	2,25E-06	0,006155	0	0,000603	8,15E-05	3,37E-06	0,000246	0,000591	0,000174	0	0,001313	5,29E-05	0,001191	4,5E-06	5,62E-07	0	0,000186	2,81E-06
о	0,007415	0	0,001095	2,02E-05	1,12E-06	4,95E-05	0	0,000784	5,62E-07	3,94E-06	0	0,67E-05	7,76E-05	2,08E-05	0,001427	0,003419	0,000229	0,000109	4,33E-05	0,002474	0,000174	1,12E-06	2,7E-05	2,36E-05	5,62E-06	
п	0,004041	1,12E-06	0,010195	1,12E-05	1,07E-05	0,000129	0,001119	0,000895	1,86E-05	1,19E-05	0,007249	0	0,000208	5,08E-06	1,12E-06	0,002649	0,000853	8,44E-06	8,1E-05	0,000551	0,000497	0,000214	1,01E-05	6,19E-06	0,000471	0,000197
р	0,002107	0	4,5E-06	0,003128	0,007484	0,004518	0,004288	0,002196	0,001605	0,000851	0,000791	0,003358	0,001594	0,005652	0,004986	0,006862	0,001608	0,001054	0,000818	0,006554	0,006762	6,3E-05	0,000758	8,77E-05	0,001306	0,001
с	4,27E-05	0	0,001236	0	5,62E-07	0	0,002482	0	0,000779	0	0	0	5,62E-05	0,00063	0	5,57E-05	0,000404	5,06E-05	0,000591	8,44E-06	3,77E-05	0,000634	2,59E-05	0	1,18E-05	3,49E-05
т	0,001354	0	0,007882	0,000138	0,000277	0,000287	0,000368	0,006108	0,00035	3,77E-05	0,004926	0	0,000249	7,25E-05	0,00022	0,000638	0,006436	4,69E-05	1,69E-05	0,000238	0,00062	0,002705	2,36E-05	0,00112	7,37E-05	8,77E-05
у	0,002951	7,87E-06	0,001465	4,89E-05	0,001784	1,07E-05	0,000307	0,003618	1,41E-05	4,27E-05	0,001392	0	0,004387	0,00235	0,000986	0,000769	0,002619	0,001698	0,000253	0,000687	0,010776	0,000761	1,24E-05	0,000157	2,7E-05	0,000394
ф	0,004394	0	0,005653	2,25E-05	0,002378	7,31E-06	0,000161	0,004276	1,69E-06	6,75E-06	0,003263	0	0,000367	0,000279	2,92E-05	0,000956	0,01385	9,11E-05	0,000226	0,000819	9,73E-05	0,001464	7,31E-06	1,24E-05	0,000128	0,000173
х	0,006272	0	8,21E-05	0,000553	0,000816	0,00117	0,001762	0,001605	0,001337	0,000699	1,24E-05	6,92E-05	0,000761	0,001312	0,000994	0,000264	2,92E-05	0,000611	0,000631	0,001213	0,001384	2,81E-06	8,44E-06	0,000355	9E-06	0,000766
ц	0,000175	0	0,000148	0	0	5,62E-07	0,000146	0	0	0,000525	0	3,94E-06	7,25E-05	1,12E-06	5,62E-07	5E-05	0	0,000382	9E-06	1,12E-05	8,44E-05	3,94E-06	0	0	5,62E-07	0
ч	0,003221	0	0,000916	0	0,000124	0	5,62E-07	3,15E-05	0	0,000163	0	0	7,25E-05	4,44E-05	0,000117	0,000215	5,62E-07	8,55E-05	4,84E-05	2,25E-05	0,000128	0	4,5E-06	0	0	6,19E-06
ш	0,000278	0	0,000569	0	1,69E-05	8,44E-06	1,12E-06	0,0009	0	0,000216	0	0,000142	2,25E-06	0	9E-06	0,000311	0	0	2,53E-06	0,000408	0	0	5,62E-07	0	0	0
щ	0	0,00022	1,12E-06	0,002027	0	1,69E-06	0	0	0,003064	0	0,001402	0	0,00026	3,71E-05	6,19E-06	0,00047	7,25E-05	0	1,97E-05	0	0,003683	0,000609	0	0	0	0,0001
ъ	9,84E-05	0	0,001391	0	1,74E-05	0	0,002208	0	0,001687	0	0,000428	0,000461	1,46E-05	0,000373	0,000299	3,66E-05	2,25E-06	0	0,001122	0,000312	5,62E-07	0	2,25E-06	0	2,81E-06	0
ы	7,31E-06	0	0,000326	0	0	0	0,001236	0	0	0,000675	0	0	0	0	0,394E-05	2,25E-06	0	1,69E-06	0	0	7,25E-05	0	0	0	0	0
ь	2,25E-06	0	0	0	0	0	0,000165	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
э	0,004067	0	0,000435	0,001046	0,000112	0,000124	0,001038	2,36E-05	5,12E-05	1,29E-05	0,001511	0,000191	0,002559	0,001372	0,000265	5,62E-07	0,001123	0,000283	0,000729	0,000574	6,19E-06	0	0,000957	7,31E-06	0,000199	0,0004
ю	0,010186	5,62E-07	9E-06	9,9E-05	6,75E-06	3,71E-05	5,01E-05	0,0012	0	0,000124	0,000102	0	0,001039	0	0,000308	0,001021	6,02E-05	0,00126	5,62E-07	0,000878	7,37E-05	0	2,25E-06	0	7,93E-05	4,95E-05
я	1,135E-05	0	0	5,62E-07	5,62E-06	2,81E-06	0	0	5,06E-06	0	6,75E-06	4,22E-05	8,6E-05	6,75E-06	3,26E-05	0	1,41E-05	2,64E-05	4,78E-05	0,000238	0	1,12E-06	3,37E-06	5,62E-07	0	5,62E-06
а	0,00312	0	5,62E-07	0,000466	0,169E-06	0,000333	1,12E-06	5,06E-06	2,53E-05	1,12E-06	0	1,07E-05	5,57E-05	2,87E-05	4,89E-05	0	0	0,000112	0,000412	0	5,62E-07	9E-06	3,04E-05	5,4E-05	0,0001	0,0004
б	0,101838	0	0	2,92E-05	0,000235	0,00017	0,000618	7,59E-05	0,000485	0,001276	1,12E-05	3,15E-05	0,000121	0,000891	0,000376	0,000597	0	8,44E-05	4,27E-05	0,000746	0,001388	5,62E-07	5,62E-07	0,00016	3,94E-05	0,000143
с	3,970323	Redundant	0,219687																							

Я	0.010838	0	0
Entropy	3.970323	Redundancy	0.219587



# Таблиця з частотами біграм з перекриттям 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	ъ	ь		
а	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	0	0		
б	0	0	0.134E-06	2.69E-06	6.72E-07	0	0	0	0	0	6.72E-07	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		
в	0	0	0.000331	0.00157	0.005495	0.001468	0.003243	0.001603	0.001434	0.004791	0.001531	0.000906	0.006531	0.010925	0.004002	0.007659	0.001482	0.002769	0.003771	0.005958	0.006599	0.005005	0.000636	0.001294	0.000153	0.001558	0.000249	0		
г	0	0	0.001258	1.08E-05	9.41E-05	1.41E-05	0.002456	4.7E-06	2.69E-06	0.0009	0	0	0.000279	0.000925	4.97E-05	0.000356	0.002655	1.41E-05	0.001322	0.000118	1.01E-05	0.001382	0	0.77E-05	8.06E-06	3.29E-05	6.05E-06	0.000208	0	
д	0	0	6.72E-07	0.000593	0.000263	0.0004	0.000374	0.000666	0.005292	6.05E-05	0.000689	0.003887	6.72E-07	0.000783	0.001278	0.000409	0.001892	0.000896	0.001066	0.000896	0.004578	0.000714	0.000004	3.16E-05	9.41E-05	4.43E-05	0.000226	0.001305	7.39E-06	1
е	0	0	0.001075	3.76E-05	0.000136	1.75E-05	0.001222	0.000628	5.38E-06	5.58E-05	0.000894	0	0.00014	0.001278	2.69E-05	0.000371	0.010734	0.000104	0.001403	0.000175	2.82E-05	0.000822	3.60E-06	6.72E-07	4.64E-05	6.05E-06	0	0		
ж	0	0	0.00494	9.27E-05	0.001155	6.05E-05	8.47E-05	0.005273	1.68E-05	4.43E-05	0.002396	6.72E-07	0.000352	0.000859	0.000034	0.002018	0.004334	0.000215	0.002561	0.0005	0.000374	0.001948	1.61E-05	6.25E-05	0.000181	7.19E-05	0.00014	0	0	
з	0	0	0.000218	0.002404	0.003467	0.000419	0.003906	0.00228	0.001406	0.002137	0.001496	0.002919	0.002132	0.001788	0.000587	0.001875	0.003091	0.000828	0.000537	0.005139	0.000626	6.72E-05	0.001404	0.000367	0.001687	0.000889	0.000764	0	0	
и	0	0	0.001535	9E-05	2.08E-05	2.35E-05	0.000811	0.004361	1.61E-05	7.39E-06	0.001471	0	0.000116	0.00176	2.08E-05	0.001379	8.8E-05	2.22E-05	4.7E-06	2.96E-05	2.15E-05	0.001919	0	3.36E-06	6.72E-07	7.26E-05	0	0		
й	0	0	0.000626	0.000205	0.000949	0.000522	0.000909	0.000374	0.000172	0.000104	0.000421	0	0.000289	0.000262	0.000349	0.002024	0.000718	0.00021	0.000265	0.000325	9E-05	0.000329	6.05E-06	4.7E-06	1.14E-05	4.3E-05	2.62E-05	1.34E-06	1	
к	0	0	0.000353	0.001379	0.005127	0.001236	0.002851	0.003334	0.000556	0.002667	0.002221	0.001545	0.003476	0.005978	0.003969	0.005728	0.002065	0.002635	0.001258	0.00554	0.000478	0.000604	0.000139	0.002	0.001364	0.00205	0.000569	0.000142	6	
л	0	0	0.000163	0.000542	0.000734	0.000356	0.000784	0.000137	0.000173	0.000202	0.000391	1.34E-06	0.000787	0.000266	0.000442	0.001138	0.000566	0.000867	0.000382	0.001709	0.000607	0.000257	6.65E-05	7.26E-05	7.32E-05	0.00033	0.00019	7.39E-06	0	0
м	0	6.72E-07	0.008415	0.000542	0.000728	0.001657	0.002276	0.000631	0.000231	0.00107	0.000195	0	0.000419	0.000504	0.000216	0.002347	0.010309	0.000406	0.00199	0.000636	0.000781	0.001802	2.15E-05	4.97E-05	2.62E-05	0.000181	2.42E-05	2.69E-06	0	0
н	0	6.72E-07	0.008173	0.00035	0.000888	0.000395	0.000695	0.005042	0.000463	0.000188	0.007916	0	0.001555	0.000197	0.000163	0.001148	0.000308	0.001054	0.000257	0.002446	0.000389	0.001597	2.75E-05	4.57E-05	9.41E-06	0.000866	5.98E-05	7.39E-06	6	0
о	0	0	0.000339	0.000404	0.000801	0.000353	0.000419	0.00346	0.000134	0.000189	0.004477	0	0.000656	0.00287	0.000273	0.002347	0.004758	0.001082	0.000257	0.001041	0.000317	0.003184	5.31E-05	8.33E-05	4.43E-05	0.00049	6.65E-05	9.41E-06	0	0
п	0	0	0.000239	0.005259	0.011566	0.000628	0.006218	0.003288	0.002376	0.001618	0.002427	0.004277	0.003187	0.000241	0.006803	0.010827	0.02473	0.003541	0.007332	0.01099	0.00927	0.000695	0.000511	0.000871	0.000445	0.002707	0.001479	0.000167	0	0
р	0	0	0.001478	2.02E-06	2.02E-06	6.72E-07	2.02E-06	0.00298	6.72E-07	0	0.000941	0	8.33E-05	0.000753	2.02E-06	6.92E-05	0.010043	6.32E-05	0.007122	1.48E-05	4.77E-05	0.000758	3.16E-05	1.34E-06	1.41E-05	4.37E-05	1.01E-05	3.36E-06	0	0
с	0	9.41E-06	0.001479	0.000204	0.002351	0.000361	0.00054	0.000387	0.000101	0.00112	0.001771	6.72E-07	0.005497	0.002897	0.001286	0.000337	0.003303	0.000234	0.000387	0.001182	0.01318	0.001075	3.9E-05	0.000201	4.03E-05	0.000519	0.000125	6.72E-07	2	0
т	0	0	0.000574	0.000288	0.003339	0.000151	0.000428	0.005255	9.07E-05	0.000118	0.004264	6.72E-07	0.000721	0.000137	0.000093	0.001629	0.016901	0.000527	0.000328	0.001487	0.000378	0.001856	3.43E-05	5.38E-05	0.000159	0.000423	1.61E-05	2.55E-05	4	0
у	0	0	0.000238	0.000904	0.001584	0.000198	0.002403	0.00329	0.001675	0.000956	0.00074	8.33E-05	0.001482	0.001678	0.001503	0.001156	0.000487	0.001385	0.000921	0.002107	0.001937	0.00013	4.9E-05	0.000486	2.35E-05	0.001357	0.000926	0.000263	0	0
ф	0	0	0.000187	4.03E-06	2.35E-05	4.03E-06	5.38E-06	0.00178	1.34E-06	1.14E-05	0.000658	0	1.55E-05	8.94E-05	5.38E-06	1.68E-05	8.33E-05	2.22E-05	0.000468	2.69E-05	2.08E-05	0.001018	4.7E-06	7.62E-07	0	2.69E-06	1.34E-06	0	0	
х	0	0	0.001141	0.000151	0.000406	0.000137	0.000189	0.001008	4.3E-05	7.59E-05	0.000535	0	0.000241	0.000243	0.000196	0.000418	0.002396	0.000438	0.000219	0.000474	0.000142	0.00024	1.81E-05	2.35E-05	8.06E-06	7.93E-05	3.16E-05	2.69E-06	2	0
ц	0	0	0.000686	8.06E-06	5.44E-05	2.08E-05	1.55E-05	0.001084	6.72E-07	8.73E-06	0.000286	0	0.000191	5.38E-06	1.34E-05	3.7E-05	0.000392	4.57E-05	6.05E-06	3.36E-05	1.68E-05	0.000762	7.32E-07	1.34E-06	1.41E-05	4.37E-05	1.01E-05	3.36E-06	0	0
ч	0	1.34E-06	0.002428	1.21E-05	2.89E-05	6.05E-06	1.21E-05	0.003661	0	6.05E-06	0.001691	0	0.00032	4.77E-05	1.14E-05	0.000591	0.000105	3.36E-05	3.56E-05	3.02E-05	0.004408	0.000377	6.72E-07	2.69E-06	0	9.41E-06	0.000148	0	0	
ш	0	0	0.001667	1.14E-05	2.82E-05	4.03E-06	8.06E-06	0.002638	1.34E-06	6.72E-06	0.002023	0	0.000519	0.000562	2.49E-05	0.000458	0.000364	5.51E-05	6.05E-06	6.72E-06	0.001046	0.000375	2.02E-06	6.72E-07	2.69E-06	3.36E-06	3.36E-06	0	0	
щ	0	0	0.00039	0	0	1.34E-06	0.001476	0	6.72E-07	0.000808	0	6.72E-07	0	6.72E-07	0	4.7E-05	2.69E-06	6.72E-07	3.36E-06	1.34E-06	6.72E-07	8.67E-05	0	0	0	0	0	0	0	
ъ	0	0	0	0	0	6.72E-07	0	0	0	0.00197	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
ь	0	0	5.71E-05	0.000782	0.001066	0.000241	0.000345	0.001355	6.65E-05	0.000224	0.000487	0.001805	0.0005	0.003111	0.001086	0.000854	0.000348	0.000514	0.000439	0.001293	0.000878	0.000114	2.02E-05	0.001187	1.55E-05	0.000326	0.000595	6.72E-06	0	0
э	0	6.72E-07	0.000652	0.000057	0.001303	0.000318	0.000463	0.001961	0.0001	0.000427	0.001076	6.72E-07	0.002006	0.000176	0.000677	0.00236	0.000898	0.001031	0.000219	0.002148	0.00056	0.000233	3.56E-05	0.000114	0.00012	0.000507	0.000536	8.73E-06	0	0
я	0	0	0	0	6.72E-07	0.00176	6.72E-06	4.7E-06	0	6.05E-06	1.34E-06	8.06E-06	5.11E-05	0.000103	8.73E-06	4.1E-05	0	1.75E-05	3.23E-05	5.71E-05	0.002845	0	1.34E-06	4.03E-06	6.72E-07	0	6.72E-07	0	0	0
я	0	0	0.000149	0.000667	0.000301	0.000128	0.000546	8.8E-05	8.13E-05	0.000104	0.000344	0	0.000292	0.000136	0.000171	0.000341	0.000249	0.000318	0.000193	0.000449	0.000615	8.06E-05	3.23E-05	3.02E-05	6.64E-05	0.000224	0.000146	0.000515	0	0
я	0	0	0.000356	0.000521	0.001566	0.000538	0.00127	0.000457	0.000703	0.001852	0.001077	7.39E-05	0.001072	0.001234	0.000764	0.0002339	0.000852	0.001096	0.000292	0.002097	0.002247	0.000266	5.58E-05	0.000263	6.52E-05	0.000529	6.11E-05	0.000247	0	0

# Таблиця з частотами біграм без перекриття 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	ъ	ь	
а	0	0.002742	0.00742	0.014716	0.004818	0.007221	0.004226	0.001941	0.004013	0.011392	7.87E-06	0.010271	0.002266	0.005638	0.0155508	0.012223	0.015397	0.003965	0.015508	0.006758	0.003903	0.0007	0.001215	0.000335	0.005543	0.000825	5.4E-05		
б	9E-06	0	0	0.12E-06	1.12E-06	0	0	0	0	0	0	0	0	0	0	0	0.12E-06	0	0.12E-06	0	0	0	0	0	0	0	0		
в	0.016304	0	7.87E-06	0.000619	0.003166	0.000793	0.001978	0.000848	0.001098	0.003724	0.000139	0.000749	0.004374	0.000876	0.002734	0.004753	5.62E-06	0.000854	0.002795	0.00345	0.004894	0.000109	0.000438	0.001009	8.44E-05	0.000733	0.00117	0.000223	
г	0.002037	0	0.01064	0	6.52E-05	9E-06	1.57E-05	0.02055	1.12E-06	0	0.00721	0	0.00222	0.000773	3.6E-05	0.000753	0.00218	0	0.00114	9.11E-05	4.5E-06	0.001169	0	6.19E-05	9E-06	1.57E-05	0.37E-06	4.5E-06	
д	0.006091	1.12E-06	0.005394	1.24E-05	3.37E-05	2.81E-05	0.00258	0.02401	0	0.000475	0.003032	0	0.001023	0.000965	0.0001	0.001176	0.002718	0.00282	0.005066	0.003196	0.001082	0.000732	1.12E-06	5.17E-05	9E-06	4.95E-05	0.01064	4.5E-06	
е	0.007443	0	0.000881	0	5.51E-05	0	0.000547	0.0005027	0	0.000679	0	8.44E-05	0.001458	0.235E-06	0.000222	0.008774	0	0.001091	5.51E-05	6.75E-06	0.000641	0	0	2.59E-05	4.5E-06	0	0		
ж	0.000429	0	0.004137	3.15E-05	0.000601	2.25E-05	3.37E-05	0.02451	1.12E-06	3.37E-05	0.000713	1.12E-06	0.00021	0.000713	7.75E-05	0.000155	0.003580	6.64E-05	0.000212	0.000319	0.000168	4.5E-06	5.06E-05	0.00014	2.91E-05	0.30E-06	0	0	
з	0.016954	0	0.04E-05	0.001054	0.002399	0.003473	0.002444	0.001568	0.00029	0.001207	0.001134	0.002372	0.000997	0.006206	0.00299	0.007024	0.0004	0.000046	0.000353	0.00033	0.0029E-05	0.001045	0.000283	0.000901	0.000645	0.000645	0.000645	0.000645	
и	0.0002	0	0.001295	7.31E-05	0	1.8E-05	0.000676	0.003686	1.46E-05	0	0.001205	0	8.77E-05	4.5E-06	4.5E-06	0.001182	6.07E-05	0	1.12E-06	1.57E-05	0	0.001614	0	0	6.41E-05	0	0	0	
й	0.001115	0	0.004944	0.000137	0.000733	0.000389	0.000686	0.000278	0.000512	6.86E-05	0.000317	0	0.00121	0.000191	0.000242	0.001651	0.000557	3.04E-05	0.000182	0.000121	1.46E-05	0.000241	0	0	5.62E-06	1.24E-05	9E-06	0	
к	0.01855	0	8.1E-05	0.000407	0.002379	0.000483	0.001569	0.002334	0.000309	0.001744	0.000719	0.001268	0.00181	0.004724	0.002766	0.0002978	0.000186	0.000178	0.000594	0.002738	0.003207	3.37E-06	3.26E-05	0.001579	0.001134	0.001172	0.000373	0.00011	
л	0.007942	0	0.135E-05	0	0	0.000151	1.12E-06	0	0	0	0	0	6.07E-05	9.34E-05	0.00025	0.000391	1.01E-05	1.12E-06	0.00195	0.000626	0.000282	0	2.25E-06	0	3.71E-05	5.29E-05	8.89E-05	0	
м	0.004782	1.12E-06	0.006994	0	0.000268	0	4.5E-06	0.000467	1.57E-05	1.24E-05	0.002381	0	2.59E-05	0.000357	3.87E-06	0.00182	0.008308	2.25E-06	0.01543	0.000117	0.000405	0.001498	0	0	7.87E-06	2.25E-05	5.62E-06	0	
н	0.00787	0	0.00673	2.36E-05	5.62E-06	0.000114	0.00024	0.003781	0.000337	4.5E-06	0.001656	0	0.000584	8.66E-05	3.7E-06	0.000261	0.00562	0.00017	0	0.001352	6.75E-05	0.001093	7.87E-06	1.12E-06	0	0	0.000173	1.62E-06	1.12E-06
о	0.007344	0	0.002673	2.02E-05	1.12E-06	0.006E-05	0.002798	0	4.5E-06	0.003079	0	5.85E-05	7.42E-05	1.69E-05	0.001128	0.003336	0.000235	5.17E-05	0.000109	3.49E-05	0.002415	1.01E-05	1.12E-06	0.00015	2.92E-05	6.75E-06	4.5E-06	0	
п	0.004629	1.12E-06	0	0.01013	1.01E-05	0.000119	0.001115	1.46E-05	6.75E-05	0.002733	0.000158	0	0.000584	0.000576	0.000529	0.000768	0.000182	0.00105	0.005454	0.000585	0.000745	6.63E-05	0.000349	0.000552	9E-06	0.001325	0.001149	0.0001	
р	0.0217	0	0.62E-05	0.003148	0.007493	0.005465	0.004263	0.002222	0.000358	0.000804	0.000719	0.003573	0.001594	0.005676	0.000529	0.000768	0.000182	0.00105	0.005454	0.000585	0.000745	6.63E-05	0.000349	0.000552	9E-06	0.001325	0.001149	0.0001	
с	4.39E-05	0	0.001248	0	0.112E-06	0	0.002489	0	0	0.000732	0	5.85E-05	0.000547	0	0.000585	0.000372	5.29E-05	0.000576	9E-06	4.61E-05	0.000631	2.02E-05	0	0	1.01E-05	3.49E-05	9E-06	4.5E-06	
т	0.001369	0	0.007888	0.000129	0.0003	0.000298	0.000378	0.000605	0.000529	3.82E-05	0.004963	0	0.000256	7.54E-05	0.000227	0.000338	0.000558	4.5E-05	0.000253	0.000616	0.000261	2.14E-05	0.000127	7.76E-06	8.66E-05	0.000219	4.39E-05	0	
у	0.002896	6.75E-06	0.001417	5.06E-05	0.001706	7.87E-06	0.000321	0.003569	6.75E-06	4.27E-05	0.001355	0	0.004482	0.002344	0.001027	0.00081	0.002571	0.01723	0.00242	0.000679	0.01095	0.000759	1.57E-05	0.000162	2.92E-05	0.000418	7.2E-05	0	
ф	0.00444	0	0.005515	2.14E-05	0.002399	6.75E-06	0.000136	0.002484	2.25E-06	7.87E-06	0.003274	0	0.003682	0.002297	2.36E-05	0.000945	0.013891	9.11E-05	0.002612	0.000793	0.000102	0.001477	5.62E-06	9E-06	0.000125	0.000117	1.12E-06	1.8E-06	
х	0.00637	0	7.99E-05	0.000574	0.000847	0.001172	0.001756	0.000155	0.001324	0.000693	9E-06	6.3E-05	0.000746	0.001327	0.001018	0.000256	2.59E-05	0.000622	0.000635	0.001243	0.001422	0.0025E-06	1.12E-05	0.00035	1.01E-05	0.000768	0.00074	0.000223	
ц	0.00189	0	0.000146	0	0	0.12E-06	0.00141	0	0	0.00052	0	1.12E-06	6.41E-05	0	0	6.07E-05	0	0.000385	4.5E-06	1.24E-05	9.45E-05	0	3.37E-06	0	0	0	0	0	
ч	0.003173	0	0.000802	0	0.00011	0	1.12E-06	3.15E-05	0	0.000153	0	0	6.97E-05	4.95E-05	0.000136	0.002251	1.12E-06	7.2E-05	5.74E-05	2.47E-05	0.000188	0	3.37E-06	0	0	7.87E-06	0		
ш	0.000529	0	0.000253	1.57E-05	9E-06	0.000122	0	0.00022	0	0.000139	2.25E-06	0	0.000139	2.25E-06	0	0	0.000139	2.25E-06	0	0	0.000139	2.25E-06	0	0	0	0	0	0	
щ	0.000216	1.12E-06	0.002099	0	2.25E-06	0	0	0.003013	0	0.001376	0	0.000247	3.94E-05	0.000465	7.99E-05	0	0	0	0.003643	0.000625	0	0	0	0	0	0	0.00011	0	
ъ	9.22E-05	0	0.001416	0	2.25E-05	0	0	0.002282	0	0	0.00171	0	0.000415	0.004457	1.12E-05	0.000382	0.000301	3.82E-05	1.12E-06	0	0.000127	0.000332	1.12E-06	0	0	3.37E-06	0	0	
ы	5.62E-06	0	0.000325	0	0	0	0.001228	0	0	0.000676	0	0	0	0	0	4.5E-05	2.25E-06	0	3.37E-06	0	0	0	6.97E-05	0	0	0	0	0	0
ь	2.25E-06	0	0	0	0	0.000157	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
а	0.004042	0	0.000445	0.001045	0.000108	0.000119	0.001019	2.7E-05	4.5E-05	1.24E-05	0.001515	0.0002	0.002538	0.001408	0.000254	1.12E-06	0.000112	0.000276	0.000747	0.000548	4.5E-06	0	0.000931	5.62E-06	0.000197	0.000444	5.62E-06	9E-06	
б	0.001215	0	6.75E-06	9.56E-05	9E-06	3.94E-05	6.07E-05	0.001226	0	0.000132	0.000106	0.00102	0	0.000331	0.000985	6.52E-05	0	0	0.000899	7.31E-05	0	2.25E-06	0	7.87E-05	4.84E-05	0.000406	9E-06	0	
в	1.12E-05	0	0	0	0	6.75E-06	4.5E-06	0	0	7.87E-06	0	6.75E-06	4.61E-05	7.99E-05	9E-06	3.04E-05	0	1.46E-05	2.25E-05	5.06E-05	0.002378	0	2.25E-06	3.37E-06	1.12E-06	0	0	0	
г	0.003414	0	0	0.000409	0	0	0.000339	1.12E-06	1.12E-06	3.04E-05	2.25E-06	0	1.01E-05	5.29E-05	3.26E-05	4.61E-05	0	0	0.006065	9.79E-05	0.000411	0	0	7.87E-06	3.71E-05	5.06E-05	0.000107	0.000414	
д	0.001036	0	0	0.259E-05	0.000231	0.000163	0.000628	8.66E-05	0.000466	0.001293	1.24E-05	2.81E-05	0.000114	0.000899	0.000216	0.000598	0	2.92E-05	4.39E-05	0.000748	0.0014	1.12E-06	1.12E-06	0.000168	3.71E-05	0.000134	1.12E-05	0.000231	
е	3.970333	Redundant	0.219587																										





$$H^{(20)}$$

Лабораторная работа №1

Произвольная часть текста:  
дую\_понравившуюся\_ж

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:  
 $1.86909513113414 < H < 2.61723922676608$

Двоичная таблица угаданных символов:

10000000000000000000000000000000
10000000000000000000000000000000
10000000000000000000000000000000
01000000000000000000000000000000
10000000000000000000000000000000

Вероятности:

$q[1] = 0.52$
$q[2] = 0.1$
$q[3] = 0.08$
$q[4] = 0.06$
$q[5] = 0.02$
$q[6] = 0.02$
$q[7] = 0.02$
$q[8] = 0$
$q[9] = 0.02$
$q[10] = 0.02$
$q[11] = 0$
$q[12] = 0$
$q[13] = 0.02$
$q[14] = 0$
$q[15] = 0.02$
$q[16] = 0$
$q[17] = 0$
$q[18] = 0$
$q[19] = 0$
$q[20] = 0$
$q[21] = 0$
$q[22] = 0$
$q[23] = 0$
$q[24] = 0.02$
$q[25] = 0$
$q[26] = 0$
$q[27] = 0$
$q[28] = 0$
$q[29] = 0$
$q[30] = 0.02$
$q[31] = 0.06$
$q[32] = 0$

Строка состояния:

$$1.86909513113414 < H^{(20)} < 2.61723922676608$$

$$0.47655215464 < R^{(20)} < 0.62618097377$$

$$H^{(30)}$$

Лабораторная работа №1

Произвольная часть текста:  
в\_большей\_мере\_зависит\_от\_чь

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:  
 $1.46317759524885 < H < 2.29243127287761$

Двоичная таблица угаданных символов:

10000000000000000000000000000000
10000000000000000000000000000000
00000100000000000000000000000000
10000000000000000000000000000000
01000000000000000000000000000000

Вероятности:

$q[1] = 0.6$
$q[2] = 0.06$
$q[3] = 0.1$
$q[4] = 0.02$
$q[5] = 0.04$
$q[6] = 0.04$
$q[7] = 0.02$
$q[8] = 0.02$
$q[9] = 0$
$q[10] = 0$
$q[11] = 0.02$
$q[12] = 0$
$q[13] = 0.02$
$q[14] = 0.02$
$q[15] = 0$
$q[16] = 0.02$
$q[17] = 0$
$q[18] = 0$
$q[19] = 0$
$q[20] = 0$
$q[21] = 0$
$q[22] = 0$
$q[23] = 0$
$q[24] = 0$
$q[25] = 0$
$q[26] = 0.02$
$q[27] = 0$
$q[28] = 0$
$q[29] = 0$
$q[30] = 0$
$q[31] = 0$
$q[32] = 0$

Строка состояния:

$$1.46317759524885 < H^{(30)} < 2.29243127287761$$

$$0.54151374542 < R^{(30)} < 0.70736448095$$

## Висновки

Під час та після виконання роботи, ми зрозуміли що являє собою поняття «ентропія» в інформаційному сенсі (чим більше ентропія  $H_n$  – тим складніше вгадати блок інформації, яку утримує в собі  $n$ -грама) і навчилися її обраховувати. Також було засвоєно поняття надлишковості, що означає собою передбачуваність мови, тобто показник  $R$  являє собою наскільки відсотків деяка мовна є передбачуваною, тобто людина чи комп'ютер з шансом  $R$  може вгадати наступну літеру даної мови.

Також, за підсумками роботи можна заявити, що відсутність пробілу хоч і зменшує алфавіт мови, але піднімає ентропію, а значить і знижує надлишковість. Також, в моєму випадку, можна побачити дуже високу наближеність ентропії і надлишковості при відсутності і наявності перекриття в біграмах, що може свідчити про стабільну структуру тексту, тобто текст має не так багато унікальних біграм.