



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
тексту

Виконали:

студенти III курсу ФТІ
групи ФБ-81
Столярчук Владислав,
Шаруєв Олександр.

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.

Постановка задачі:

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 $(10) H$, $(20) H$, $(30) H$.
3. Використовуючи отримані значення ентропії, оцінити надлишковість російської мови в різних моделях джерела.

Хід роботи:

- 1) Уважно прочитали методичні вказівки до виконання комп'ютерного практикуму.
- 2) Скопіювали в файл prest.txt текст російською мовою довжини 1.1 Мб. Текст взято з роману “Злочин і кара” Ф.М. Достоевського.
- 3) Записали текст з файлу prest.txt в один рядок (string).
- 4) Текст треба було підготувати до роботи, тому була зроблена така фільтрація:
 - видалено символи, що не є текстовими;
 - прописні букви перетворено у малі літери;
 - декілька пробілів поспіль перетворені в один пробіл.
- 5) Додатково записали текст без пробілів.
- 6) Відповідно до завдання знайшли частоту монограм в тексті з пробілами та без.

Заповнили таблицю:

Таблиця частот монограм у тексті з пробілами(зліва) та без пробілів (справа)

—	0.17404	о	0.11427
о	0.09438	е	0.08705
е	0.07190	а	0.08035
а	0.06637	т	0.06519
т	0.05385	н	0.06461
н	0.05337	и	0.06438
и	0.05318	с	0.05295
с	0.04374	в	0.04648
в	0.03839	л	0.04596
л	0.03796	р	0.04184
р	0.03456	к	0.03283
к	0.02712	д	0.03193
д	0.02637	м	0.03131
м	0.02586	у	0.02952
у	0.02438	п	0.02741
п	0.02264	ь	0.02344
ь	0.01936	я	0.02180
я	0.01800	ч	0.01833
ч	0.01514	б	0.01731
б	0.01430	г	0.01707
г	0.01410	ы	0.01641
ы	0.01356	з	0.01530
з	0.01264	ж	0.01140
ж	0.00942	й	0.00996
й	0.00823	х	0.00847

х	0.00700	ш	0.00830
ш	0.00685	ю	0.00560
ю	0.00463	э	0.00352
э	0.00291	щ	0.00299
щ	0.00247	ц	0.00278
ц	0.00230	ф	0.00123
ф	0.00101		

7) Підраховали частоти біграм з перетинами та без перетинів, з текстом без пробілів та з ними.

З перетинами біграми будуть братись так: “абвгд” аб, бв, вг, гд.

Без перетинів біграми будуть братись так: “абвгд” аб, вг.

Таблиця частот біграм з перетинами у тексті з пробілами:

а 0.00251	т 0.00880	ае 0.00117	ач 0.00090	бо 0.00189
б 0.00680	у 0.00453	аж 0.00159	аш 0.00088	бр 0.00114
в 0.01664	ф 0.00039	аз 0.00333	ащ 0.00024	бс 0.00019
г 0.00312	х 0.00154	аи 0.00011	аэ 0.00000	бт 0.00000
д 0.00803	ц 0.00029	ай 0.00088	аю 0.00084	бу 0.00125
е 0.00409	ч 0.00636	ак 0.00546	ая 0.00223	бх 0.00004
ж 0.00226	ш 0.00064	ал 0.00699	б 0.00044	бц 0.00000
з 0.00438	щ 0.00007	ам 0.00334	ба 0.00061	бч 0.00002
и 0.01145	ь 0.00000	ан 0.00309	бб 0.00000	бш 0.00001
й 0.00001	э 0.00288	ао 0.00002	бв 0.00005	бщ 0.00016
к 0.00800	ю 0.00005	ап 0.00075	бд 0.00002	бы 0.00366
л 0.00243	я 0.00263	ар 0.00167	бе 0.00210	бь 0.00018
м 0.00576	а 0.01705	ас 0.00478	бж 0.00000	бю 0.00001
н 0.01578	аа 0.00003	ат 0.00426	би 0.00079	бя 0.00060
о 0.01139	аб 0.00049	ау 0.00006	бк 0.00013	в 0.00555
п 0.01583	ав 0.00277	аф 0.00004	бл 0.00068	ва 0.00584
р 0.00493	аг 0.00064	ах 0.00091	бм 0.00004	вб 0.00000
с 0.01554	ад 0.00180	ац 0.00004	бн 0.00025	вв 0.00002

вг 0.00001	гт 0.00001	ег 0.00325	жр 0.00001	ид 0.00167
вд 0.00084	гу 0.00044	ед 0.00266	жс 0.00001	ие 0.00164
ве 0.00495	гх 0.00000	ее 0.00172	жт 0.00000	иж 0.00021
вз 0.00053	гц 0.00000	еж 0.00078	жу 0.00026	из 0.00158
ви 0.00310	гч 0.00004	ез 0.00116	жц 0.00000	ии 0.00052
вй 0.00000	гш 0.00000	еи 0.00010	жч 0.00000	ий 0.00094
вк 0.00011	гъ 0.00000	ей 0.00209	жы 0.00000	ик 0.00228
вл 0.00051	гя 0.00000	ек 0.00113	жь 0.00005	ил 0.00481
vm 0.00010	д 0.00086	ел 0.00545	жю 0.00000	им 0.00244
вн 0.00151	да 0.00509	ем 0.00415	з 0.00119	ин 0.00297
во 0.00610	дб 0.00003	ен 0.00635	за 0.00442	ио 0.00020
вп 0.00038	дв 0.00089	ео 0.00022	зб 0.00011	ип 0.00022
вр 0.00049	дг 0.00001	еп 0.00118	зв 0.00084	ир 0.00078
вс 0.00360	дд 0.00005	ер 0.00584	зг 0.00034	ис 0.00187
вт 0.00029	де 0.00416	ес 0.00431	зд 0.00067	ит 0.00415
ву 0.00058	дж 0.00002	ет 0.00573	зе 0.00017	иу 0.00001
vx 0.00006	дз 0.00001	еу 0.00016	зж 0.00005	иф 0.00001
вц 0.00002	ди 0.00229	еф 0.00001	зз 0.00003	их 0.00165
вч 0.00019	дк 0.00022	ех 0.00051	зи 0.00034	иц 0.00080
вш 0.00061	дл 0.00056	ец 0.00032	зй 0.00000	ич 0.00148
вщ 0.00000	дм 0.00007	еч 0.00116	зк 0.00011	иш 0.00061
вы 0.00265	дн 0.00178	еш 0.00100	зл 0.00023	ищ 0.00013
вь 0.00016	до 0.00360	ещ 0.00093	зм 0.00019	иы 0.00000
вю 0.00000	дп 0.00008	еы 0.00000	зн 0.00174	иь 0.00000
вя 0.00018	др 0.00183	ез 0.00000	зо 0.00048	изэ 0.00000
г 0.00105	дс 0.00026	ею 0.00023	зп 0.00001	ию 0.00030
га 0.00097	дт 0.00027	ея 0.00021	зр 0.00023	ия 0.00107
гв 0.00001	ду 0.00195	ж 0.00075	зс 0.00001	й 0.00592
гг 0.00000	дж 0.00004	жа 0.00120	зт 0.00001	йа 0.00000
гд 0.00109	дц 0.00026	жб 0.00002	зу 0.00060	йб 0.00002
ге 0.00017	дч 0.00005	жг 0.00000	зч 0.00002	йг 0.00000
гж 0.00001	дш 0.00007	жд 0.00069	зш 0.00001	йд 0.00027
гз 0.00000	дщ 0.00000	же 0.00433	зы 0.00026	йе 0.00000
ги 0.00050	ды 0.00042	жж 0.00001	зь 0.00014	йз 0.00000
гк 0.00005	дъ 0.00107	жи 0.00123	зю 0.00000	йи 0.00000
гл 0.00145	дю 0.00001	жй 0.00000	зя 0.00042	йк 0.00013
гм 0.00004	дя 0.00042	жк 0.00010	и 0.01723	йл 0.00020
гн 0.00027	е 0.01884	жл 0.00002	иа 0.00007	йм 0.00004
го 0.00745	еа 0.00001	жм 0.00001	иб 0.00065	йн 0.00034
гр 0.00051	еб 0.00130	жн 0.00069	ив 0.00241	йо 0.00000
гс 0.00002	ев 0.00108	жо 0.00002	иг 0.00052	йп 0.00000

йс	0.00042	ли	0.00526	мш	0.00001	ож	0.00206	пъ	0.00015
йт	0.00046	лк	0.00026	мы	0.00080	оз	0.00089	пю	0.00000
йу	0.00000	лл	0.00005	мь	0.00008	ои	0.00061	пя	0.00063
йф	0.00000	лм	0.00001	мю	0.00000	ой	0.00277	р	0.00070
йх	0.00000	лн	0.00025	мя	0.00034	ок	0.00155	ра	0.00730
йц	0.00003	ло	0.00523	н	0.00422	ол	0.00569	рб	0.00011
йч	0.00018	лп	0.00006	на	0.00987	ом	0.00505	рв	0.00041
йш	0.00018	лр	0.00000	нб	0.00001	он	0.00645	рг	0.00014
йщ	0.00000	лс	0.00141	нв	0.00001	оо	0.00022	рд	0.00030
к	0.00474	лт	0.00011	нг	0.00002	оп	0.00140	ре	0.00507
ка	0.00674	лу	0.00144	нд	0.00028	ор	0.00487	рж	0.00023
кб	0.00000	лф	0.00000	не	0.00998	ос	0.00559	рз	0.00005
кв	0.00024	лч	0.00021	нж	0.00000	от	0.00670	ри	0.00514
кд	0.00000	лш	0.00000	нз	0.00001	оу	0.00009	рй	0.00000
ке	0.00037	лщ	0.00000	ни	0.00740	оф	0.00017	рк	0.00019
кж	0.00001	лы	0.00060	нк	0.00025	ох	0.00046	рл	0.00006
кэ	0.00001	ль	0.00458	нм	0.00000	оц	0.00008	рм	0.00028
ки	0.00194	лю	0.00073	нн	0.00237	оч	0.00223	рн	0.00062
кй	0.00000	ля	0.00122	но	0.00949	ош	0.00098	ро	0.00700
кк	0.00001	м	0.00734	нп	0.00000	ощ	0.00022	рп	0.00010
кл	0.00050	ма	0.00286	нр	0.00005	оэ	0.00000	рр	0.00002
км	0.00000	мб	0.00001	нс	0.00025	ою	0.00057	рс	0.00011
кн	0.00055	мв	0.00000	нт	0.00036	оя	0.00061	рт	0.00067
ко	0.00848	мг	0.00008	ну	0.00276	п	0.00004	ру	0.00250
кп	0.00000	мд	0.00000	нф	0.00003	па	0.00078	рф	0.00029
кр	0.00143	ме	0.00385	нх	0.00000	пг	0.00000	рх	0.00006
кс	0.00023	ми	0.00260	нц	0.00017	пе	0.00285	рц	0.00002
кт	0.00063	мй	0.00000	нч	0.00019	пи	0.00078	рч	0.00007
ку	0.00117	мк	0.00007	нш	0.00000	пк	0.00007	рш	0.00021
кх	0.00003	мл	0.00009	нщ	0.00008	пл	0.00061	рщ	0.00001
кц	0.00000	мм	0.00002	ны	0.00242	пн	0.00009	ры	0.00080
кш	0.00001	мн	0.00156	нь	0.00118	по	0.00892	рь	0.00094
л	0.00689	мо	0.00363	ню	0.00021	пп	0.00003	рю	0.00020
ла	0.00554	мп	0.00004	ня	0.00176	пр	0.00657	ря	0.00096
лб	0.00003	мр	0.00006	о	0.02351	пс	0.00003	с	0.00363
лв	0.00001	мс	0.00010	оа	0.00001	пт	0.00007	са	0.00166
лг	0.00013	мт	0.00016	об	0.00339	пу	0.00074	сб	0.00009
лд	0.00002	му	0.00211	ов	0.00768	пц	0.00001	св	0.00133
ле	0.00363	мф	0.00001	ог	0.00403	пч	0.00002	сг	0.00003
лж	0.00030	мц	0.00001	од	0.00480	пш	0.00001	сд	0.00020
лз	0.00001	мч	0.00005	ое	0.00168	пы	0.00022	се	0.00440

сж	0.00002	тр	0.00303	уэ	0.00000	цв	0.00002	шл	0.00052
сз	0.00001	тс	0.00106	ую	0.00071	це	0.00073	шм	0.00001
си	0.00138	тт	0.00008	уя	0.00006	ци	0.00011	шн	0.00039
сй	0.00000	ту	0.00161	ф	0.00003	цк	0.00002	шо	0.00018
ск	0.00343	тх	0.00001	фа	0.00021	цл	0.00000	шп	0.00000
сл	0.00263	тц	0.00006	фе	0.00005	цм	0.00000	шр	0.00000
см	0.00109	тч	0.00032	фи	0.00029	цо	0.00019	шс	0.00000
сн	0.00087	тш	0.00001	фк	0.00000	цр	0.00000	шт	0.00004
со	0.00267	тщ	0.00003	фл	0.00001	цт	0.00001	шу	0.00020
сп	0.00167	ты	0.00137	фм	0.00000	цу	0.00018	шч	0.00000
ср	0.00015	ть	0.00678	фо	0.00010	цч	0.00000	шш	0.00000
сс	0.00059	тэ	0.00001	фп	0.00000	цы	0.00015	шь	0.00062
ст	0.00953	тю	0.00008	фр	0.00006	ч	0.00046	шю	0.00000
су	0.00068	тя	0.00039	фс	0.00000	ча	0.00267	щ	0.00001
сф	0.00000	у	0.00677	фт	0.00001	чв	0.00000	ща	0.00036
сх	0.00018	уа	0.00003	фу	0.00010	че	0.00384	щг	0.00000
сц	0.00005	уб	0.00071	фф	0.00001	чз	0.00000	ще	0.00134
сч	0.00036	ув	0.00064	фч	0.00000	чи	0.00152	щи	0.00052
сш	0.00007	уг	0.00142	фш	0.00000	чк	0.00050	щл	0.00000
сщ	0.00000	уд	0.00193	фы	0.00002	чл	0.00001	щн	0.00005
сы	0.00019	уе	0.00018	фь	0.00010	чм	0.00000	щр	0.00000
сь	0.00304	уж	0.00174	х	0.00268	чн	0.00082	щу	0.00015
сю	0.00025	уз	0.00022	ха	0.00040	чо	0.00008	щь	0.00004
ся	0.00352	уи	0.00001	хв	0.00018	чр	0.00008	ы	0.00433
т	0.00576	уй	0.00022	хе	0.00032	чт	0.00411	ыа	0.00000
та	0.00579	ук	0.00063	хж	0.00000	чу	0.00066	ыб	0.00015
тб	0.00001	ул	0.00144	хз	0.00001	чч	0.00000	ыв	0.00075
тв	0.00210	ум	0.00155	хи	0.00051	чш	0.00014	ыг	0.00011
тг	0.00001	ун	0.00051	хк	0.00000	чы	0.00000	ыд	0.00016
тд	0.00008	уо	0.00000	хл	0.00014	чь	0.00024	ые	0.00069
те	0.00575	уп	0.00064	хм	0.00007	ш	0.00008	ыж	0.00002
тж	0.00000	ур	0.00040	хн	0.00015	ша	0.00088	ыз	0.00005
тз	0.00001	ус	0.00099	хо	0.00221	шв	0.00001	ыи	0.00001
ти	0.00320	ут	0.00155	хр	0.00008	шт	0.00000	ый	0.00121
тй	0.00000	уу	0.00000	хс	0.00002	шд	0.00000	ык	0.00015
тк	0.00033	уф	0.00001	хт	0.00005	ше	0.00215	ыл	0.00186
тл	0.00020	ух	0.00039	ху	0.00014	шж	0.00000	ым	0.00085
тм	0.00003	уц	0.00000	хш	0.00002	шз	0.00000	ын	0.00016
тн	0.00107	уч	0.00078	хэ	0.00000	ши	0.00133	ыо	0.00000
то	0.01459	уш	0.00056	ц	0.00037	шй	0.00000	ып	0.00011
тп	0.00007	ущ	0.00027	ца	0.00050	шк	0.00044	ыр	0.00019

ыс 0.00062	ьо 0.00000	эл 0.00000	юк 0.00001	яз 0.00019
ыт 0.00074	ьп 0.00001	эм 0.00000	юл 0.00001	яи 0.00002
ыу 0.00000	ьр 0.00000	эн 0.00002	юм 0.00005	яй 0.00009
ых 0.00077	ьс 0.00084	эп 0.00001	юн 0.00004	як 0.00017
ыц 0.00000	ьт 0.00026	эр 0.00000	юп 0.00000	ял 0.00081
ыч 0.00017	ьу 0.00000	эс 0.00001	юр 0.00005	ям 0.00032
ыш 0.00045	ьф 0.00001	эт 0.00272	юс 0.00021	ян 0.00058
ыщ 0.00001	ьх 0.00012	эф 0.00001	ют 0.00035	яо 0.00000
ыя 0.00000	ьц 0.00006	эх 0.00004	юх 0.00001	яп 0.00005
ь 0.01205	ьч 0.00008	ээ 0.00001	юч 0.00010	яр 0.00007
ьб 0.00007	ьш 0.00029	эю 0.00000	юш 0.00006	яс 0.00065
ьв 0.00004	ьщ 0.00003	ю 0.00294	ющ 0.00016	ят 0.00170
ьг 0.00013	ьь 0.00000	юб 0.00030	юю 0.00004	ях 0.00016
ьд 0.00002	ью 0.00031	юв 0.00001	я 0.01169	яц 0.00007
ье 0.00043	ья 0.00067	юг 0.00000	яа 0.00000	яч 0.00014
ьз 0.00014	э 0.00001	юд 0.00028	яб 0.00001	яш 0.00001
ьи 0.00012	эв 0.00000	юе 0.00000	яв 0.00024	ящ 0.00012
ьк 0.00128	эг 0.00001	юж 0.00001	яг 0.00007	яь 0.00000
ьл 0.00000	эд 0.00001	юз 0.00001	яд 0.00048	яю 0.00009
ьм 0.00035	эй 0.00001	юи 0.00000	яе 0.00008	яя 0.000
ьн 0.00203	эк 0.00004	юй 0.00000	яж 0.00010	

Таблиця частот біграм з перетинами у тексті без пробілів:

аа 0.00026	ас 0.00635	бд 0.00002	бч 0.00004	вй 0.00000
аб 0.00106	ат 0.00524	бе 0.00211	бш 0.00001	вк 0.00053
ав 0.00449	ау 0.00052	бж 0.00000	бщ 0.00016	вл 0.00059
аг 0.00093	аф 0.00008	бз 0.00001	бы 0.00366	вм 0.00021
ад 0.00260	ах 0.00104	би 0.00082	бь 0.00018	вн 0.00203
ае 0.00165	ац 0.00007	бк 0.00014	бэ 0.00009	во 0.00645
аж 0.00175	ач 0.00157	бл 0.00068	бю 0.00001	вп 0.00091
аз 0.00375	аш 0.00096	бм 0.00004	бя 0.00063	вр 0.00068
аи 0.00138	ащ 0.00025	бн 0.00027	ва 0.00591	вс 0.00423
ай 0.00088	аь 0.00000	бо 0.00198	вб 0.00018	вт 0.00070
ак 0.00633	аэ 0.00028	бп 0.00001	вв 0.00039	ву 0.00075
ал 0.00725	аю 0.00085	бр 0.00114	вг 0.00023	вф 0.00001
ам 0.00397	ая 0.00252	бс 0.00021	вд 0.00110	вх 0.00009
ан 0.00472	ба 0.00062	бт 0.00002	ве 0.00509	вц 0.00003
ао 0.00113	бб 0.00000	бу 0.00129	вж 0.00004	вч 0.00033
ап 0.00236	бв 0.00007	бх 0.00005	вз 0.00065	вш 0.00064
ар 0.00212	бг 0.00000	бц 0.00000	ви 0.00332	вщ 0.00000

вы 0.00265	ди 0.00234	ет 0.00665	зб 0.00013	ил 0.00502
вь 0.00016	дй 0.00000	еу 0.00067	зв 0.00096	им 0.00296
вэ 0.00026	дк 0.00028	еф 0.00005	зг 0.00037	ин 0.00471
вю 0.00000	дл 0.00059	ех 0.00077	зд 0.00076	ио 0.00138
вя 0.00022	дм 0.00009	ец 0.00035	зе 0.00020	ип 0.00194
га 0.00099	дн 0.00191	еч 0.00171	зж 0.00005	ир 0.00128
гб 0.00002	до 0.00364	еш 0.00107	зз 0.00005	ис 0.00368
гв 0.00012	дп 0.00015	ещ 0.00094	зи 0.00037	ит 0.00506
гг 0.00002	др 0.00186	еы 0.00000	зй 0.00000	иу 0.00051
гд 0.00114	дс 0.00035	ез 0.00038	зк 0.00020	иф 0.00006
ге 0.00019	дт 0.00030	ею 0.00023	зл 0.00025	их 0.00183
гж 0.00001	ду 0.00196	ея 0.00044	зм 0.00025	иц 0.00084
гз 0.00005	дф 0.00000	жа 0.00121	зн 0.00186	ич 0.00203
ги 0.00056	дх 0.00005	жб 0.00005	зо 0.00054	иш 0.00067
гк 0.00010	дц 0.00026	жв 0.00005	зп 0.00015	ищ 0.00013
гл 0.00146	дч 0.00008	жг 0.00001	зр 0.00026	иы 0.00000
гм 0.00006	дш 0.00007	жд 0.00072	зс 0.00012	иь 0.00000
гн 0.00036	дщ 0.00000	же 0.00434	зт 0.00010	из 0.00020
го 0.00757	ды 0.00042	жж 0.00001	зу 0.00064	ию 0.00030
гп 0.00014	дь 0.00107	жз 0.00001	зф 0.00000	ия 0.00133
гр 0.00057	дэ 0.00002	жи 0.00130	зх 0.00001	йа 0.00008
гс 0.00012	дю 0.00001	жй 0.00000	зц 0.00001	йб 0.00020
гт 0.00004	дя 0.00043	жк 0.00018	зч 0.00005	йв 0.00047
гу 0.00048	еа 0.00023	жл 0.00003	зш 0.00001	йт 0.00020
гф 0.00000	еб 0.00215	жм 0.00002	зщ 0.00000	йд 0.00059
гх 0.00001	ев 0.00309	жн 0.00079	зы 0.00026	йе 0.00008
гц 0.00000	ег 0.00362	жо 0.00009	зь 0.00014	йж 0.00009
гч 0.00005	ед 0.00360	жп 0.00005	зэ 0.00002	йз 0.00013
гш 0.00000	ее 0.00208	жр 0.00002	зю 0.00000	йи 0.00050
гъ 0.00000	еж 0.00097	жс 0.00006	зя 0.00042	йк 0.00044
гэ 0.00001	ез 0.00184	жт 0.00005	иа 0.00027	йл 0.00029
гю 0.00000	ей 0.00133	жу 0.00027	иб 0.00137	йм 0.00031
гя 0.00001	ей 0.00209	жх 0.00000	ив 0.00421	йн 0.00081
да 0.00511	ек 0.00180	жц 0.00000	иг 0.00083	йо 0.00031
дб 0.00005	ел 0.00579	жч 0.00002	ид 0.00256	йп 0.00061
дв 0.00097	ем 0.00500	жы 0.00000	ие 0.00201	йр 0.00029
дг 0.00002	ен 0.00786	жь 0.00005	иж 0.00038	йс 0.00103
дд 0.00010	ео 0.00140	жэ 0.00004	из 0.00207	йт 0.00067
де 0.00418	еп 0.00320	жю 0.00000	ии 0.00144	йу 0.00013
дж 0.00002	ер 0.00638	жя 0.00003	ий 0.00094	йф 0.00003
дз 0.00005	ес 0.00620	за 0.00444	ик 0.00315	йх 0.00005

йц	0.00005	ле	0.00395	мр	0.00028	нь	0.00118	пи	0.00078
йч	0.00048	лж	0.00033	мс	0.00088	нэ	0.00004	пк	0.00008
йш	0.00023	лз	0.00017	мт	0.00043	ню	0.00021	пл	0.00061
йщ	0.00000	ли	0.00576	му	0.00235	ня	0.00180	пм	0.00000
йэ	0.00008	лк	0.00057	мф	0.00004	оа	0.00021	пн	0.00010
йю	0.00001	лл	0.00013	мх	0.00007	об	0.00464	по	0.00893
йя	0.00006	лм	0.00014	мц	0.00002	ов	0.01033	пп	0.00003
ка	0.00679	лн	0.00085	мч	0.00033	ог	0.00445	пр	0.00657
кб	0.00051	ло	0.00617	мш	0.00004	од	0.00588	пс	0.00003
кв	0.00066	лп	0.00059	мщ	0.00000	ое	0.00235	пт	0.00007
кг	0.00008	лр	0.00037	мы	0.00080	ож	0.00254	пу	0.00074
кд	0.00022	лс	0.00196	мь	0.00008	оз	0.00149	пх	0.00000
ке	0.00045	лт	0.00038	мэ	0.00010	ои	0.00191	пц	0.00001
кж	0.00012	лу	0.00156	мю	0.00000	ой	0.00277	пч	0.00002
кз	0.00009	лф	0.00002	мя	0.00046	ок	0.00254	пш	0.00001
ки	0.00228	лх	0.00003	на	0.00992	ол	0.00604	пы	0.00022
кй	0.00000	лц	0.00000	нб	0.00022	ом	0.00590	пь	0.00015
кк	0.00022	лч	0.00055	нв	0.00052	он	0.00851	пю	0.00000
кл	0.00058	лш	0.00002	нг	0.00009	оо	0.00190	пя	0.00063
км	0.00017	лщ	0.00000	нд	0.00043	оп	0.00349	ра	0.00732
кн	0.00099	лы	0.00060	не	0.01006	ор	0.00543	рб	0.00014
ко	0.00877	ль	0.00458	нж	0.00002	ос	0.00770	рв	0.00047
кп	0.00032	лэ	0.00009	нз	0.00013	от	0.00801	рг	0.00016
кр	0.00158	лю	0.00073	ни	0.00767	оу	0.00080	рд	0.00032
кс	0.00063	ля	0.00134	нк	0.00044	оф	0.00022	ре	0.00508
кт	0.00088	ма	0.00296	нл	0.00005	ох	0.00066	рж	0.00023
ку	0.00128	мб	0.00024	нм	0.00011	оц	0.00012	рз	0.00006
кф	0.00001	мв	0.00069	нн	0.00273	оч	0.00318	ри	0.00521
кх	0.00007	мг	0.00024	но	0.00976	ош	0.00104	рй	0.00000
кц	0.00001	мд	0.00040	нп	0.00047	ощ	0.00023	рк	0.00022
кч	0.00019	ме	0.00397	нр	0.00025	оз	0.00036	рл	0.00007
кш	0.00002	мж	0.00014	нс	0.00071	ою	0.00058	рм	0.00028
кщ	0.00000	мз	0.00020	нт	0.00056	оя	0.00109	рн	0.00069
кэ	0.00009	ми	0.00316	ну	0.00291	па	0.00078	ро	0.00704
кю	0.00000	мй	0.00000	нф	0.00004	пб	0.00001	рп	0.00025
кя	0.00010	мк	0.00038	нх	0.00004	пв	0.00000	рр	0.00004
ла	0.00568	мл	0.00018	нц	0.00017	пг	0.00000	рс	0.00016
лб	0.00024	мм	0.00022	нч	0.00031	пд	0.00000	рт	0.00070
лв	0.00063	мн	0.00222	нш	0.00002	пе	0.00285	ру	0.00251
лг	0.00025	мо	0.00418	нщ	0.00008	пж	0.00000	рф	0.00029
лд	0.00027	мп	0.00080	ны	0.00242	пз	0.00000	рх	0.00006

рц 0.00002	тб 0.00035	ул 0.00152	хг 0.00006	цс 0.00003
рч 0.00009	тв 0.00263	ум 0.00186	хд 0.00018	цт 0.00003
рш 0.00021	тг 0.00011	ун 0.00128	хе 0.00037	цу 0.00019
рщ 0.00001	тд 0.00037	уо 0.00039	хж 0.00004	цф 0.00000
ры 0.00080	те 0.00591	уп 0.00123	хз 0.00007	цх 0.00000
рь 0.00094	тж 0.00011	ур 0.00057	хи 0.00073	цч 0.00003
рэ 0.00001	тз 0.00013	ус 0.00153	хк 0.00014	цш 0.00000
рю 0.00020	ти 0.00364	ут 0.00192	хл 0.00020	цы 0.00015
ря 0.00097	тй 0.00000	уу 0.00013	хм 0.00015	цэ 0.00001
са 0.00172	тк 0.00061	уф 0.00003	хн 0.00036	цю 0.00000
сб 0.00022	тл 0.00026	ух 0.00045	хо 0.00237	ця 0.00001
св 0.00162	тм 0.00022	уц 0.00001	хп 0.00033	ча 0.00267
сг 0.00008	тн 0.00159	уч 0.00118	хр 0.00015	чб 0.00002
сд 0.00039	то 0.01496	уш 0.00058	хс 0.00028	чв 0.00006
се 0.00446	тп 0.00053	ущ 0.00028	хт 0.00017	чг 0.00001
сж 0.00022	тр 0.00318	уз 0.00009	ху 0.00022	чд 0.00002
сз 0.00008	тс 0.00152	ую 0.00072	хф 0.00001	че 0.00384
си 0.00153	тт 0.00039	уя 0.00020	хх 0.00001	чж 0.00000
сй 0.00000	ту 0.00176	фа 0.00022	хц 0.00001	чз 0.00001
ск 0.00363	тф 0.00001	фб 0.00000	хч 0.00008	чи 0.00155
сл 0.00270	тх 0.00005	фв 0.00000	хш 0.00003	чк 0.00053
см 0.00125	тц 0.00007	фд 0.00000	хщ 0.00000	чл 0.00002
сн 0.00128	тч 0.00058	фе 0.00005	хэ 0.00005	чм 0.00001
со 0.00287	тш 0.00003	фи 0.00029	хю 0.00000	чн 0.00087
сп 0.00202	тщ 0.00003	фк 0.00000	хя 0.00003	чо 0.00011
ср 0.00022	ты 0.00137	фл 0.00001	ца 0.00051	чп 0.00005
сс 0.00089	ть 0.00678	фм 0.00001	цб 0.00001	чр 0.00011
ст 0.00981	тэ 0.00011	фо 0.00010	цв 0.00006	чс 0.00003
су 0.00084	тю 0.00008	фп 0.00001	цг 0.00002	чт 0.00413
сф 0.00001	тя 0.00048	фр 0.00006	цд 0.00001	чу 0.00067
сх 0.00022	уа 0.00019	фс 0.00000	це 0.00074	чх 0.00001
сц 0.00006	уб 0.00088	фт 0.00001	цж 0.00000	чч 0.00001
сч 0.00045	ув 0.00131	фу 0.00010	цз 0.00001	чш 0.00014
сш 0.00009	уг 0.00153	фф 0.00001	ци 0.00014	чы 0.00000
сщ 0.00000	уд 0.00226	фч 0.00000	цк 0.00003	чь 0.00024
сы 0.00019	уе 0.00029	фш 0.00000	цл 0.00000	чэ 0.00000
сь 0.00304	уж 0.00185	фы 0.00002	цм 0.00001	чя 0.00001
сэ 0.00006	уз 0.00038	фь 0.00010	цн 0.00003	ша 0.00089
сю 0.00025	уй 0.00055	ха 0.00044	цо 0.00022	шб 0.00000
ся 0.00355	уй 0.00022	хб 0.00011	цп 0.00005	шв 0.00001
та 0.00596	ук 0.00095	хв 0.00042	цр 0.00001	шт 0.00000

шд 0.00000	ья 0.00005	ьж 0.00010	эс 0.00001	юя 0.00005
ше 0.00215	ыб 0.00032	ьз 0.00043	эт 0.00272	яа 0.00030
шж 0.00000	ыв 0.00119	ьи 0.00117	эф 0.00001	яб 0.00038
шз 0.00000	ыг 0.00018	ьй 0.00000	эх 0.00004	яв 0.00151
ши 0.00133	ыд 0.00038	ьк 0.00193	эч 0.00000	яг 0.00026
шй 0.00000	ые 0.00081	ьл 0.00018	ээ 0.00001	яд 0.00107
шк 0.00044	ыж 0.00006	ьм 0.00070	эю 0.00000	яе 0.00038
шл 0.00052	ыз 0.00018	ьн 0.00338	юа 0.00007	яж 0.00025
шм 0.00001	ыи 0.00039	ьо 0.00079	юб 0.00037	яз 0.00049
шн 0.00040	ый 0.00121	ьп 0.00110	юв 0.00025	яи 0.00096
шо 0.00018	ык 0.00033	ьр 0.00026	юг 0.00006	яй 0.00009
шп 0.00001	ыл 0.00191	ьс 0.00180	юд 0.00043	як 0.00067
шр 0.00001	ым 0.00106	ьт 0.00089	юе 0.00005	ял 0.00096
шс 0.00001	ын 0.00061	ьу 0.00026	юж 0.00007	ям 0.00064
шт 0.00005	ыо 0.00026	ьф 0.00004	юз 0.00008	ян 0.00183
шу 0.00020	ып 0.00044	ьх 0.00026	юи 0.00024	яо 0.00071
шф 0.00000	ыр 0.00029	ьц 0.00009	юй 0.00000	яп 0.00109
шц 0.00000	ыс 0.00101	ьч 0.00067	юк 0.00022	яр 0.00045
шч 0.00000	ыт 0.00103	ьш 0.00034	юл 0.00005	яс 0.00172
шш 0.00000	ыу 0.00013	ьщ 0.00003	юм 0.00016	ят 0.00234
шь 0.00062	ыф 0.00000	ьь 0.00000	юн 0.00027	яу 0.00029
шю 0.00000	ых 0.00080	ьэ 0.00020	юо 0.00017	яф 0.00003
шя 0.00000	ыц 0.00001	ью 0.00031	юп 0.00029	ях 0.00029
ша 0.00036	ыч 0.00029	ья 0.00092	юр 0.00015	яц 0.00009
шб 0.00000	ыш 0.00046	эв 0.00001	юс 0.00047	яч 0.00055
шт 0.00000	ыщ 0.00001	эг 0.00001	ют 0.00049	яш 0.00006
шд 0.00000	ыз 0.00009	эд 0.00001	юу 0.00007	ящ 0.00012
ше 0.00134	ыю 0.00000	эи 0.00000	юф 0.00002	яь 0.00000
ши 0.00052	ыя 0.00005	эй 0.00001	юх 0.00003	яэ 0.00016
шл 0.00000	ба 0.00024	эк 0.00004	юц 0.00001	яю 0.00009
шн 0.00005	ьб 0.00043	эл 0.00000	юч 0.00024	яя 0.00022
шр 0.00000	ьв 0.00122	эм 0.00000	юш 0.00007	
шу 0.00015	ьг 0.00028	эн 0.00002	ющ 0.00017	
шч 0.00000	ьд 0.00053	эп 0.00001	юэ 0.00004	
шь 0.00004	ье 0.00083	эр 0.00000	юю 0.00004	

Таблиця частот біграм без перетинів у тексті з пробілами:

а	0.00248	аз	0.00330	бх	0.00005	гг	0.00000	дц	0.00025
б	0.00686	аи	0.00011	бц	0.00000	гд	0.00107	дч	0.00005
в	0.01672	ай	0.00087	бч	0.00002	ге	0.00018	дш	0.00007
г	0.00309	ак	0.00545	бш	0.00001	гж	0.00000	дщ	0.00000
д	0.00816	ал	0.00702	бщ	0.00016	гз	0.00000	ды	0.00042
е	0.00412	ам	0.00334	бы	0.00369	ги	0.00048	дь	0.00109
ж	0.00225	ан	0.00312	бь	0.00018	гк	0.00005	дю	0.00001
з	0.00438	ао	0.00002	бю	0.00001	гл	0.00148	дя	0.00041
и	0.01146	ап	0.00078	бя	0.00060	гм	0.00004	е	0.01876
й	0.00001	ар	0.00167	в	0.00549	гн	0.00029	еа	0.00001
к	0.00807	ас	0.00471	ва	0.00584	го	0.00746	еб	0.00127
л	0.00233	ат	0.00417	вв	0.00004	гр	0.00052	ев	0.00101
м	0.00577	ау	0.00006	вг	0.00001	гс	0.00002	ег	0.00326
н	0.01577	аф	0.00004	вд	0.00077	гт	0.00002	ед	0.00268
о	0.01151	ах	0.00093	ве	0.00497	гу	0.00044	ее	0.00175
п	0.01573	ац	0.00004	вз	0.00058	гч	0.00005	еж	0.00073
р	0.00491	ач	0.00090	ви	0.00314	гь	0.00000	ез	0.00117
с	0.01564	аш	0.00091	вк	0.00012	д	0.00088	еи	0.00011
т	0.00873	ащ	0.00026	вл	0.00049	да	0.00501	ей	0.00209
у	0.00450	аз	0.00000	вм	0.00011	дб	0.00002	ек	0.00114
ф	0.00039	аю	0.00085	вн	0.00147	дв	0.00091	ел	0.00549
х	0.00149	ая	0.00223	во	0.00612	дг	0.00001	ем	0.00411
ц	0.00027	б	0.00042	вп	0.00037	дд	0.00006	ен	0.00632
ч	0.00644	ба	0.00065	вр	0.00053	де	0.00404	ео	0.00022
ш	0.00065	бв	0.00005	вс	0.00354	дж	0.00001	еп	0.00117
щ	0.00007	бд	0.00003	вт	0.00027	дз	0.00001	ер	0.00588
ь	0.00000	бе	0.00214	ву	0.00061	ди	0.00216	ес	0.00437
э	0.00277	бж	0.00000	вх	0.00006	дк	0.00021	ет	0.00565
ю	0.00005	би	0.00075	вц	0.00001	дл	0.00058	еу	0.00017
я	0.00268	бк	0.00013	вч	0.00020	дм	0.00007	еф	0.00001
а	0.01708	бл	0.00065	вш	0.00059	дн	0.00182	ех	0.00050
аа	0.00003	бм	0.00004	вщ	0.00000	до	0.00356	ец	0.00030
аб	0.00047	бн	0.00025	вы	0.00270	дп	0.00009	еч	0.00110
ав	0.00279	бо	0.00191	вь	0.00015	др	0.00181	еш	0.00099
аг	0.00064	бр	0.00120	вя	0.00017	дс	0.00023	ещ	0.00092
ад	0.00183	бс	0.00019	г	0.00098	дт	0.00024	еы	0.00000
ае	0.00117	бт	0.00000	га	0.00099	ду	0.00196	ею	0.00022
аж	0.00159	бу	0.00121	гв	0.00001	дх	0.00005	ея	0.00020

ж	0.00075	зш	0.00001	йи	0.00000	лв	0.00001	му	0.00214
жа	0.00123	зы	0.00024	йк	0.00013	лг	0.00012	мф	0.00001
жб	0.00003	зь	0.00013	йл	0.00021	лд	0.00002	мц	0.00000
жг	0.00000	зя	0.00040	йм	0.00004	ле	0.00367	мч	0.00004
жд	0.00070	и	0.01737	йн	0.00033	лж	0.00031	мш	0.00000
же	0.00424	иа	0.00006	йо	0.00001	лз	0.00001	мы	0.00082
жж	0.00001	иб	0.00064	йп	0.00000	ли	0.00522	мь	0.00007
жи	0.00125	ив	0.00239	йс	0.00039	лк	0.00029	мя	0.00031
жк	0.00011	иг	0.00049	йт	0.00045	лл	0.00005	н	0.00423
жл	0.00002	ид	0.00173	йф	0.00000	лм	0.00001	на	0.00989
жм	0.00000	ие	0.00171	йх	0.00000	лн	0.00024	нб	0.00001
жн	0.00073	иж	0.00020	йц	0.00003	ло	0.00526	нв	0.00001
жо	0.00002	из	0.00162	йч	0.00017	лп	0.00006	нг	0.00002
жр	0.00001	ии	0.00053	йш	0.00020	лр	0.00000	нд	0.00029
жс	0.00001	ий	0.00101	йщ	0.00000	лс	0.00138	не	0.01007
жт	0.00001	ик	0.00231	к	0.00469	лт	0.00013	нж	0.00000
жу	0.00028	ил	0.00474	ка	0.00671	лу	0.00147	нз	0.00001
жь	0.00005	им	0.00248	кб	0.00000	лф	0.00000	ни	0.00733
жю	0.00000	ин	0.00304	кв	0.00023	лч	0.00021	нк	0.00025
з	0.00116	ио	0.00021	кд	0.00001	лш	0.00000	нм	0.00000
за	0.00438	ип	0.00022	ке	0.00033	лщ	0.00000	нн	0.00236
зб	0.00012	ир	0.00082	кж	0.00001	лы	0.00059	но	0.00944
зв	0.00087	ис	0.00183	кз	0.00000	ль	0.00466	нр	0.00006
зг	0.00035	ит	0.00417	ки	0.00189	лю	0.00075	нс	0.00025
зд	0.00068	иу	0.00001	кй	0.00000	ля	0.00119	нт	0.00037
зе	0.00018	иф	0.00001	кк	0.00001	м	0.00731	ну	0.00270
зж	0.00005	их	0.00165	кл	0.00053	ма	0.00282	нф	0.00004
зз	0.00002	иц	0.00081	км	0.00001	мб	0.00001	нх	0.00000
зи	0.00035	ич	0.00148	кн	0.00050	мг	0.00006	нц	0.00018
зй	0.00000	иш	0.00061	ко	0.00848	мд	0.00000	нч	0.00017
зк	0.00011	ищ	0.00012	кп	0.00000	ме	0.00381	нш	0.00000
зл	0.00024	иь	0.00000	кр	0.00141	ми	0.00256	нщ	0.00009
зм	0.00017	ию	0.00029	кс	0.00021	мк	0.00007	ны	0.00234
зн	0.00171	ия	0.00107	кт	0.00063	мл	0.00007	нь	0.00119
зо	0.00049	й	0.00596	ку	0.00112	мм	0.00003	ню	0.00021
зп	0.00002	йа	0.00000	кх	0.00003	мн	0.00161	ня	0.00176
зр	0.00023	йб	0.00003	кц	0.00000	мо	0.00363	о	0.02346
зс	0.00002	йг	0.00000	кш	0.00001	мп	0.00004	оа	0.00001
зт	0.00001	йд	0.00026	л	0.00701	мр	0.00007	об	0.00333
зу	0.00057	йе	0.00000	ла	0.00555	мс	0.00010	ов	0.00769
зч	0.00001	йз	0.00000	лб	0.00002	мт	0.00016	ог	0.00408

од	0.00483	пш	0.00002	се	0.00440	тс	0.00105	ую	0.00071
ое	0.00170	пы	0.00022	сж	0.00003	тт	0.00008	уя	0.00005
ож	0.00213	пь	0.00015	сз	0.00001	ту	0.00157	ф	0.00002
оз	0.00093	пю	0.00000	си	0.00131	тх	0.00002	фа	0.00022
ои	0.00064	пя	0.00063	ск	0.00341	тц	0.00005	фе	0.00006
ой	0.00268	р	0.00069	сл	0.00255	тч	0.00034	фи	0.00028
ок	0.00155	ра	0.00736	см	0.00110	тш	0.00001	фк	0.00000
ол	0.00561	рб	0.00012	сн	0.00088	тщ	0.00003	фл	0.00001
ом	0.00508	рв	0.00040	со	0.00261	ты	0.00145	фм	0.00001
он	0.00638	рг	0.00015	сп	0.00173	ть	0.00683	фо	0.00011
оо	0.00025	рд	0.00031	ср	0.00014	тэ	0.00001	фп	0.00000
оп	0.00134	ре	0.00508	сс	0.00058	тю	0.00008	фр	0.00007
ор	0.00474	рж	0.00023	ст	0.00969	тя	0.00037	фс	0.00000
ос	0.00554	рз	0.00005	су	0.00067	у	0.00677	фт	0.00001
от	0.00671	ри	0.00516	сф	0.00000	уа	0.00003	фу	0.00009
оу	0.00010	рк	0.00020	сх	0.00020	уб	0.00073	фф	0.00001
оф	0.00017	рл	0.00006	сц	0.00005	ув	0.00065	фч	0.00000
ох	0.00047	рм	0.00026	сч	0.00038	уг	0.00148	фш	0.00000
оц	0.00008	рн	0.00063	сш	0.00006	уд	0.00198	фы	0.00002
оч	0.00218	ро	0.00705	сы	0.00019	уе	0.00018	фь	0.00009
ош	0.00098	рп	0.00009	сь	0.00309	уж	0.00170	х	0.00266
ощ	0.00022	рр	0.00002	сю	0.00021	уз	0.00022	ха	0.00040
оз	0.00000	рс	0.00012	ся	0.00355	уи	0.00001	хв	0.00017
ою	0.00055	рт	0.00066	т	0.00577	уй	0.00023	хе	0.00037
оя	0.00057	ру	0.00246	та	0.00573	ук	0.00066	хж	0.00000
п	0.00004	рф	0.00030	тб	0.00001	ул	0.00143	хз	0.00001
па	0.00082	рх	0.00006	тв	0.00204	ум	0.00162	хи	0.00048
пг	0.00000	рц	0.00002	тг	0.00001	ун	0.00052	хк	0.00000
пе	0.00286	рч	0.00006	тд	0.00009	уо	0.00000	хл	0.00013
пи	0.00075	рш	0.00022	те	0.00582	уп	0.00063	хм	0.00007
пк	0.00008	рщ	0.00002	тж	0.00000	ур	0.00041	хн	0.00016
пл	0.00057	ры	0.00081	тз	0.00001	ус	0.00102	хо	0.00220
пн	0.00011	рь	0.00092	ти	0.00314	ут	0.00156	хр	0.00008
по	0.00905	рю	0.00024	тй	0.00000	уу	0.00001	хс	0.00001
пп	0.00003	ря	0.00093	тк	0.00035	уф	0.00002	хт	0.00006
пр	0.00655	с	0.00359	тл	0.00022	ух	0.00040	ху	0.00015
пс	0.00002	са	0.00171	тм	0.00003	уц	0.00000	хш	0.00001
пт	0.00007	сб	0.00009	тн	0.00109	уч	0.00075	ц	0.00039
пу	0.00072	св	0.00137	то	0.01466	уш	0.00055	ца	0.00052
пц	0.00001	сг	0.00003	тп	0.00007	ущ	0.00027	цв	0.00002
пч	0.00002	сд	0.00019	тр	0.00297	уэ	0.00000	це	0.00076

ци	0.00011	шн	0.00038	ыр	0.00015	э	0.00001	юч	0.00011
цк	0.00002	шо	0.00017	ыс	0.00064	эв	0.00000	юш	0.00005
цл	0.00000	шп	0.00000	ыт	0.00071	эг	0.00001	ющ	0.00017
цо	0.00021	шс	0.00000	ых	0.00076	эд	0.00000	юю	0.00004
цр	0.00000	шт	0.00004	ыц	0.00000	эй	0.00001	я	0.01165
цт	0.00000	шу	0.00018	ыч	0.00016	эк	0.00004	яа	0.00000
цу	0.00019	шч	0.00000	ыш	0.00043	эл	0.00000	яб	0.00000
цы	0.00014	шш	0.00000	ыщ	0.00001	эм	0.00000	яв	0.00024
ч	0.00047	шь	0.00060	ыя	0.00000	эн	0.00002	яг	0.00007
ча	0.00269	шю	0.00000	ь	0.01195	еп	0.00002	яд	0.00051
че	0.00382	щ	0.00001	ьб	0.00007	эр	0.00000	яе	0.00009
чз	0.00000	ща	0.00034	ьв	0.00004	эс	0.00001	яж	0.00009
чи	0.00153	ше	0.00136	ьг	0.00013	эт	0.00282	яз	0.00021
чк	0.00053	щи	0.00049	ьд	0.00002	эф	0.00001	яи	0.00001
чл	0.00001	щл	0.00000	ье	0.00042	эх	0.00005	яй	0.00010
чм	0.00000	щн	0.00004	ьз	0.00014	ээ	0.00001	як	0.00019
чн	0.00086	щр	0.00000	ьи	0.00011	эю	0.00000	ял	0.00082
чо	0.00009	щу	0.00014	ьк	0.00129	ю	0.00292	ям	0.00037
чр	0.00008	щъ	0.00003	ьл	0.00000	юб	0.00027	ян	0.00060
чт	0.00408	ы	0.00426	ьм	0.00035	юв	0.00001	яп	0.00004
чу	0.00063	ьб	0.00014	ьн	0.00200	юг	0.00000	яр	0.00006
чч	0.00000	ьв	0.00078	ьо	0.00000	юд	0.00031	яс	0.00063
чш	0.00015	ьг	0.00011	ьп	0.00000	юе	0.00000	ят	0.00172
чы	0.00000	ьд	0.00016	ьр	0.00000	юж	0.00001	яx	0.00016
чъ	0.00023	ье	0.00072	ьс	0.00085	юз	0.00001	яц	0.00007
ш	0.00007	ьж	0.00002	ьт	0.00023	юи	0.00000	яч	0.00014
ша	0.00088	ьз	0.00004	ьф	0.00001	юк	0.00001	яш	0.00002
ше	0.00219	ьи	0.00001	ьх	0.00012	юл	0.00001	ящ	0.00013
шж	0.00000	ьй	0.00123	ьц	0.00006	юм	0.00004	яю	0.00010
шз	0.00000	ьк	0.00017	ьч	0.00009	юн	0.00005	яя	0.000
ши	0.00137	ьл	0.00182	ьш	0.00028	юр	0.00005		
шк	0.00045	ьм	0.00081	ьщ	0.00003	юс	0.00022		
шл	0.00052	ьн	0.00019	ью	0.00033	ют	0.00033		
шм	0.00001	ьп	0.00011	ья	0.00070	юх	0.00000		

Таблиця частот біграм без перетинів у тексті без пробілів:

аа 0.00025	бл 0.00068	вф 0.00001	да 0.00502	ез 0.00181
аб 0.00106	бм 0.00004	вх 0.00008	дб 0.00006	ей 0.00129
ав 0.00449	бн 0.00031	вц 0.00003	дв 0.00099	ей 0.00210
аг 0.00092	бо 0.00204	вч 0.00038	дг 0.00002	ек 0.00191
ад 0.00256	бп 0.00001	вш 0.00065	дд 0.00012	ел 0.00584
ае 0.00167	бр 0.00113	вщ 0.00001	де 0.00410	ем 0.00502
аж 0.00176	бс 0.00024	вы 0.00260	дж 0.00003	ен 0.00788
аз 0.00378	бт 0.00003	вь 0.00015	дз 0.00004	ео 0.00144
аи 0.00138	бу 0.00130	вэ 0.00027	ди 0.00238	еп 0.00323
ай 0.00084	бх 0.00004	вю 0.00001	дй 0.00000	ер 0.00652
ак 0.00644	бч 0.00003	вя 0.00022	дк 0.00029	ес 0.00622
ал 0.00721	бш 0.00001	га 0.00098	дл 0.00060	ет 0.00662
ам 0.00406	бщ 0.00015	гб 0.00003	дм 0.00009	еу 0.00067
ан 0.00473	бы 0.00371	гв 0.00012	дн 0.00189	еф 0.00004
ао 0.00113	бь 0.00018	гг 0.00002	до 0.00374	ех 0.00073
ап 0.00233	бэ 0.00012	гд 0.00114	дп 0.00015	ец 0.00036
ар 0.00214	бю 0.00001	ге 0.00021	др 0.00185	еч 0.00168
ас 0.00639	бя 0.00064	гж 0.00001	дс 0.00037	еш 0.00105
ат 0.00515	ва 0.00600	гз 0.00006	дт 0.00029	ещ 0.00095
ау 0.00053	вб 0.00019	ги 0.00058	ду 0.00202	еы 0.00000
аф 0.00007	вв 0.00040	гк 0.00009	дф 0.00000	еэ 0.00036
ах 0.00103	вг 0.00021	гл 0.00146	дх 0.00006	ею 0.00022
ац 0.00007	вд 0.00105	гм 0.00006	дц 0.00026	ея 0.00045
ач 0.00159	ве 0.00507	гн 0.00033	дч 0.00008	жа 0.00123
аш 0.00098	вж 0.00003	го 0.00756	дш 0.00007	жб 0.00004
ащ 0.00026	вз 0.00062	гп 0.00014	дщ 0.00000	жв 0.00005
аз 0.00028	ви 0.00331	гр 0.00054	ды 0.00042	жг 0.00001
аю 0.00092	вй 0.00000	гс 0.00014	дь 0.00107	жд 0.00070
ая 0.00253	вк 0.00056	гт 0.00003	дэ 0.00001	же 0.00434
ба 0.00063	вл 0.00064	гу 0.00048	дю 0.00001	жж 0.00001
бб 0.00000	вм 0.00021	гф 0.00001	дя 0.00046	жз 0.00002
бв 0.00007	вн 0.00201	гх 0.00000	еа 0.00024	жи 0.00131
бд 0.00002	во 0.00649	гц 0.00000	еб 0.00210	жй 0.00000
бе 0.00205	вп 0.00096	гч 0.00006	ев 0.00311	жк 0.00018
бж 0.00001	вр 0.00071	гш 0.00000	ег 0.00364	жл 0.00002
бз 0.00001	вс 0.00418	гэ 0.00001	ед 0.00356	жм 0.00002
би 0.00082	вт 0.00070	гю 0.00000	ее 0.00203	жн 0.00073
бк 0.00016	ву 0.00073	гя 0.00001	еж 0.00101	жо 0.00010

жп 0.00006	иа 0.00023	йм 0.00033	кш 0.00002	мж 0.00014
жр 0.00002	иб 0.00137	йн 0.00083	кщ 0.00001	мз 0.00021
жс 0.00005	ив 0.00413	йо 0.00030	кэ 0.00010	ми 0.00309
жт 0.00005	иг 0.00082	йп 0.00063	кю 0.00000	мк 0.00042
жу 0.00028	ид 0.00266	йр 0.00028	кя 0.00010	мл 0.00016
жх 0.00000	ие 0.00206	йс 0.00107	ла 0.00565	мм 0.00023
жч 0.00002	иж 0.00038	йт 0.00070	лб 0.00024	мн 0.00221
жъ 0.00006	из 0.00207	йу 0.00011	лв 0.00063	мо 0.00418
жэ 0.00005	ии 0.00144	йф 0.00004	лг 0.00027	мп 0.00080
жя 0.00002	ий 0.00086	йх 0.00005	лд 0.00027	мр 0.00029
за 0.00441	ик 0.00318	йц 0.00006	ле 0.00393	мс 0.00088
зб 0.00011	ил 0.00495	йч 0.00048	лж 0.00033	мт 0.00047
зв 0.00093	им 0.00296	йш 0.00023	лз 0.00017	му 0.00231
зг 0.00037	ин 0.00477	йщ 0.00000	ли 0.00584	мф 0.00004
зд 0.00073	ио 0.00138	йэ 0.00007	лк 0.00059	мх 0.00007
зе 0.00018	ип 0.00192	йю 0.00001	лл 0.00012	мц 0.00001
зж 0.00007	ир 0.00129	йя 0.00006	лм 0.00014	мч 0.00031
зз 0.00005	ис 0.00370	ка 0.00671	лн 0.00085	мш 0.00004
зи 0.00039	ит 0.00488	кб 0.00051	ло 0.00623	мщ 0.00000
зй 0.00000	иу 0.00050	кв 0.00063	лп 0.00057	мы 0.00079
зк 0.00021	иф 0.00006	кг 0.00007	лр 0.00039	мь 0.00008
зл 0.00026	их 0.00183	кд 0.00022	лс 0.00200	мэ 0.00011
зм 0.00028	иц 0.00081	ке 0.00044	лт 0.00038	мю 0.00001
зн 0.00191	ич 0.00202	кж 0.00010	лу 0.00155	мя 0.00045
зо 0.00058	иш 0.00063	кз 0.00008	лф 0.00002	на 0.00990
зп 0.00015	ищ 0.00015	ки 0.00229	лх 0.00004	нб 0.00022
зр 0.00026	иы 0.00000	кй 0.00000	лц 0.00001	нв 0.00054
зс 0.00013	из 0.00019	кк 0.00023	лч 0.00056	нг 0.00008
зт 0.00010	ию 0.00034	кл 0.00058	лш 0.00002	нд 0.00046
зу 0.00059	ия 0.00134	км 0.00016	лщ 0.00001	не 0.01018
зф 0.00000	йа 0.00009	кн 0.00089	лы 0.00061	нж 0.00002
зх 0.00001	йб 0.00020	ко 0.00873	ль 0.00463	нз 0.00013
зц 0.00001	йв 0.00046	кп 0.00031	лэ 0.00008	ни 0.00774
зч 0.00004	йг 0.00021	кр 0.00161	лю 0.00071	нк 0.00044
зш 0.00002	йд 0.00061	кс 0.00059	ля 0.00134	нл 0.00005
зщ 0.00000	йе 0.00008	кт 0.00088	ма 0.00294	нм 0.00012
зы 0.00025	йж 0.00009	ку 0.00115	мб 0.00024	нн 0.00266
зь 0.00015	йз 0.00011	кф 0.00000	мв 0.00071	но 0.01000
зэ 0.00002	йи 0.00051	кх 0.00008	мг 0.00024	нп 0.00047
зю 0.00000	йк 0.00046	кц 0.00002	мд 0.00042	нр 0.00023
зя 0.00044	йл 0.00029	кч 0.00021	ме 0.00388	нс 0.00074

нт 0.00053	оя 0.00108	рт 0.00070	сю 0.00026	уй 0.00054
ну 0.00294	па 0.00075	ру 0.00246	ся 0.00348	уй 0.00021
нф 0.00004	пб 0.00001	рф 0.00027	та 0.00606	ук 0.00095
нх 0.00005	пг 0.00001	рх 0.00006	тб 0.00037	ул 0.00151
нц 0.00017	пд 0.00000	рц 0.00001	тв 0.00266	ум 0.00190
нч 0.00028	пе 0.00285	рч 0.00009	тг 0.00010	ун 0.00126
нш 0.00002	пи 0.00079	рш 0.00019	тд 0.00038	уо 0.00035
нщ 0.00007	пк 0.00006	рщ 0.00002	те 0.00593	уп 0.00120
ны 0.00245	пл 0.00059	ры 0.00079	тж 0.00011	ур 0.00060
нь 0.00119	пн 0.00010	рь 0.00090	тз 0.00014	ус 0.00154
нэ 0.00004	по 0.00909	рэ 0.00001	ти 0.00363	ут 0.00190
ню 0.00020	пп 0.00004	рю 0.00019	тй 0.00000	уу 0.00015
ня 0.00181	пр 0.00658	ря 0.00097	тк 0.00064	уф 0.00004
оа 0.00022	пс 0.00003	са 0.00170	тл 0.00029	ух 0.00046
об 0.00455	пт 0.00006	сб 0.00021	тм 0.00026	уц 0.00001
ов 0.01040	пу 0.00073	св 0.00158	тн 0.00153	уч 0.00123
ог 0.00444	пх 0.00000	сг 0.00010	то 0.01501	уш 0.00059
од 0.00574	пц 0.00001	сд 0.00042	тп 0.00051	ущ 0.00027
ое 0.00229	пч 0.00002	се 0.00443	тр 0.00314	уэ 0.00011
ож 0.00245	пш 0.00002	сж 0.00022	тс 0.00155	ую 0.00080
оз 0.00153	пы 0.00021	сз 0.00008	тт 0.00041	уя 0.00021
ои 0.00192	пь 0.00017	си 0.00154	ту 0.00180	фа 0.00020
ой 0.00270	пя 0.00066	ск 0.00359	тф 0.00002	фб 0.00000
ок 0.00259	ра 0.00725	сл 0.00269	тх 0.00006	фв 0.00000
ол 0.00594	рб 0.00014	см 0.00129	тц 0.00008	фд 0.00001
ом 0.00587	рв 0.00045	сн 0.00127	тч 0.00057	фе 0.00005
он 0.00844	рг 0.00014	со 0.00283	тш 0.00003	фи 0.00031
оо 0.00187	рд 0.00031	сп 0.00203	тщ 0.00003	фк 0.00000
оп 0.00343	ре 0.00502	ср 0.00022	ты 0.00141	фл 0.00001
ор 0.00547	рж 0.00020	сс 0.00093	ть 0.00684	фм 0.00002
ос 0.00762	рз 0.00005	ст 0.00979	тэ 0.00013	фо 0.00008
от 0.00800	ри 0.00523	су 0.00083	тю 0.00009	фп 0.00001
оу 0.00078	рй 0.00000	сф 0.00001	тя 0.00045	фр 0.00006
оф 0.00022	рк 0.00019	сх 0.00021	уа 0.00020	фс 0.00000
ох 0.00064	рл 0.00008	сц 0.00006	уб 0.00085	фт 0.00001
оц 0.00011	рм 0.00028	сч 0.00042	ув 0.00130	фу 0.00010
оч 0.00308	рн 0.00070	сш 0.00008	уг 0.00159	фф 0.00001
ош 0.00107	ро 0.00716	сщ 0.00000	уд 0.00227	фы 0.00002
ощ 0.00023	рп 0.00027	сы 0.00021	уе 0.00032	фь 0.00011
оэ 0.00037	рр 0.00004	сь 0.00297	уж 0.00194	ха 0.00046
ою 0.00056	рс 0.00017	сэ 0.00006	уз 0.00040	хб 0.00009

хв 0.00038	цс 0.00003	ши 0.00133	ым 0.00102	ьщ 0.00004
хг 0.00006	цт 0.00002	шк 0.00048	ын 0.00060	ьь 0.00000
хд 0.00018	цу 0.00020	шл 0.00049	ьо 0.00025	ьэ 0.00021
хе 0.00036	цч 0.00003	шм 0.00001	ып 0.00043	ью 0.00033
хж 0.00005	цш 0.00000	шн 0.00034	ыр 0.00029	ья 0.00093
хз 0.00007	цы 0.00017	шо 0.00018	ыс 0.00103	эв 0.00001
хи 0.00069	цэ 0.00001	шп 0.00001	ыт 0.00104	эг 0.00001
хк 0.00014	цю 0.00000	шр 0.00000	ыу 0.00010	эд 0.00001
хл 0.00020	ча 0.00270	шс 0.00001	ыф 0.00001	эи 0.00000
хм 0.00016	чб 0.00002	шт 0.00005	ых 0.00080	эй 0.00001
хн 0.00035	чв 0.00006	шу 0.00020	ыч 0.00028	эк 0.00005
хо 0.00246	чг 0.00001	шф 0.00000	ыш 0.00047	эл 0.00000
хп 0.00030	чд 0.00002	шц 0.00000	ыщ 0.00001	эм 0.00000
хр 0.00015	че 0.00390	шч 0.00000	ыз 0.00010	эн 0.00002
хс 0.00029	чж 0.00000	шш 0.00000	ью 0.00000	еп 0.00001
хт 0.00017	чз 0.00002	шь 0.00062	ья 0.00006	эр 0.00000
ху 0.00024	чи 0.00161	шю 0.00000	ьа 0.00023	эс 0.00001
хф 0.00000	чк 0.00052	шы 0.00000	ьб 0.00042	эт 0.00263
хх 0.00001	чл 0.00003	ща 0.00032	ьв 0.00120	эф 0.00001
хц 0.00001	чм 0.00001	щб 0.00000	ьг 0.00031	эх 0.00004
хч 0.00007	чн 0.00088	щг 0.00000	ьд 0.00054	эч 0.00000
хш 0.00003	чо 0.00010	щд 0.00000	ье 0.00083	ээ 0.00001
хэ 0.00006	чп 0.00005	ще 0.00137	ьж 0.00011	юа 0.00005
хю 0.00000	чр 0.00011	щи 0.00053	ьз 0.00045	юб 0.00038
хя 0.00004	чс 0.00003	щл 0.00000	ьи 0.00117	юв 0.00025
ца 0.00053	чт 0.00416	щн 0.00005	ьй 0.00000	юг 0.00005
цб 0.00000	чу 0.00065	щр 0.00001	ьк 0.00189	юд 0.00044
цв 0.00005	чх 0.00001	щу 0.00014	ьл 0.00017	юе 0.00004
цг 0.00002	чч 0.00001	щъ 0.00004	ьм 0.00070	юж 0.00006
цд 0.00001	чш 0.00015	ья 0.00006	ьн 0.00334	юз 0.00008
це 0.00071	чы 0.00000	ьб 0.00031	ьо 0.00080	юи 0.00024
цж 0.00000	чь 0.00026	ьв 0.00116	ьп 0.00107	юй 0.00000
цз 0.00001	чэ 0.00001	ьг 0.00017	ьр 0.00026	юк 0.00021
ци 0.00013	чя 0.00001	ьд 0.00038	ьс 0.00179	юл 0.00006
цк 0.00004	ша 0.00090	ье 0.00081	ьт 0.00089	юм 0.00014
цл 0.00000	шб 0.00000	ьж 0.00005	ьу 0.00023	юн 0.00023
цм 0.00001	шв 0.00001	ьз 0.00017	ьф 0.00004	юо 0.00014
цн 0.00003	шг 0.00000	ьи 0.00040	ьх 0.00028	юп 0.00028
цо 0.00021	шд 0.00000	ьй 0.00127	ьц 0.00011	юр 0.00014
цп 0.00005	ше 0.00221	ьк 0.00037	ьч 0.00064	юс 0.00046
цр 0.00001	шз 0.00000	ьл 0.00185	ьш 0.00033	ют 0.00048

юу 0.00005	юю 0.00003	яж 0.00026	яо 0.00074	яц 0.00011
юф 0.00002	юя 0.00006	яз 0.00046	яп 0.00108	яч 0.00055
юх 0.00003	яа 0.00026	яи 0.00101	яр 0.00045	яш 0.00006
юц 0.00001	яб 0.00042	яй 0.00009	яс 0.00176	ящ 0.00011
юч 0.00023	яв 0.00153	як 0.00063	ят 0.00234	яэ 0.00016
юш 0.00009	яг 0.00026	ял 0.00099	яу 0.00030	яю 0.00010
ющ 0.00015	яд 0.00101	ям 0.00058	яф 0.00004	яя 0.00022
юэ 0.00004	яе 0.00039	ян 0.00178	ях 0.00031	

8) Обрахували значення ентропії для монограм та біграм, значення занесли в таблицю.

Текст	n	H_n
з пробілами	1	4.34187
без пробілів	1	4.44938
з пробілами	2(з перетином)	3.94664
без пробілів	2(з перетином)	4.12740
з пробілами	2(без перетину)	3.94613
без пробілів	2(без перетину)	4.12696

Значення ентропії $H^{(10)}$ знаходиться в проміжку $2,5628 < H^{(10)} < 3,2697$:

Значення ентропії $H^{(20)}$ знаходиться в проміжку $1,9899 < H^{(20)} < 2,7524$:

Значення ентропії $H^{(30)}$ знаходиться в проміжку $1,1724 < H^{(30)} < 1,8445$:

Лабораторная работа №1

Произвольная часть текста:
o_в_этом_роде_относительно_че

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:52

Поле ввода символов:

Продолжить

Другой

Неравенство для энтропии:
1.17249988489048 < H < 1.84459630375829

Двоичная таблица угаданных символов:
00000000000000000000000000000000
10000000000000000000000000000000
10000000000000000000000000000000
00001000000000000000000000000000
10000000000000000000000000000000
10000000000000000000000000000000
nnnnnnnnnn1nnnnnnnnnnnnnnnnnnnnnnnn

Вероятности:

q[1] = 0.62745098039
q[2] = 0.19607843137
q[3] = 0.01960784313
q[4] = 0.01960784313
q[5] = 0.01960784313
q[6] = 0.01960784313
q[7] = 0
q[8] = 0.01960784313
q[9] = 0
q[10] = 0.0196078431
q[11] = 0
q[12] = 0
q[13] = 0
q[14] = 0
q[15] = 0
q[16] = 0
q[17] = 0.0392156862
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
q[27] = 0
q[28] = 0
q[29] = 0.0196078431
q[30] = 0
q[31] = 0
q[32] = 0

Строка состояния:

10) Оцінили надлишковість для російської мови у різних моделях відкритого тексту.

$R = 1 - H_{\infty} / H_0$

Текст	n	R
з пробілами	1	0.13163
без пробілів	1	0.11012
з пробілами	2(з перетином)	0.21067
без пробілів	2(з перетином)	0.17452
з пробілами	2(без перетину)	0.21077
без пробілів	2(без перетину)	0.17461
з пробілами	10	0.34606<R<0.48744
з пробілами	20	0.44952<R<0.60202
з пробілами	30	0.63110<R<0.76552

Висновки:

Засвоїли поняття ентропії на символ джерела та його надлишковості, вивчили та порівняли різні моделі джерела відкритого тексту для наближеного визначення ентропії, набули практичних навичок щодо оцінки ентропії на символ джерела. Попрактикувалися з фільтрацією від пунктуаційних знаків, пробілів, переносу рядка, літер верхнього і нижнього регістру. За допомогою програми CoolPinkProgram визначили значення ентропії.