



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

КРИПТОГРАФІЯ
КОМП'ЮТЕРНИЙ ПРАКТИКУМ №1

Виконала студентка гр. ФБ-24:
Тішевська Анна

Київ–2025

Мета роботи

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

Порядок виконання роботи

0. Уважно прочитати методичні вказівки до виконання комп'ютерного практикуму. 1. Написати програми для підрахунку частот букв і частот біграм в тексті, а також підрахунку H_1 та H_2 за безпосереднім означенням. Підрахувати частоти букв та біграм, а

також значення H_1 та H_2 на довільно обраному тексті російською мовою достатньої довжини (щонайменше 1Мб), де імовірності замінити відповідними частотами. Також одержати значення H_1 та H_2 на тому ж тексті, в якому вилучено 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.

1. Теоретичні відомості

Ентропія символів (монограм):

$$H_1 = - \sum_i p_i \log_2 p_i$$

Ентропія біграм:

$$H_2 = - \frac{1}{2} \sum_{i,j} p_{ij} \log_2 p_{ij}$$

Надлишковість:

$$R = 1 - \frac{H}{\log_2 |A|}$$

де $|A|$ — кількість символів у алфавіті.

Умовна ентропія $H^{(n)}$: показує середню невизначеність наступного символу при врахуванні n попередніх символів.

2. Вихідні дані

Вхідний текст: фрагмент російського корпусу (≈ 1 МБ після нормалізації).

Алфавіт:

з пробілами: 32 символи (абвгдежзийклмнопрстуфхцчшщъыэюя);

без пробілів: 31 символ.

1 153 КБ

том лев николаевич толстой война и мир том часть первая е поместья мой верный раб ну что князь **генуа** и **лукка** стали не больше как поместьями фамилии **бонапарте** нет я вас предупреждаю если вы мне не скажете что у нас война если вы еще позволите себе защищать все гадости все ужасы этого антихриста право я верю что он антихрист я вас больше не знаю вы уж не друг мой вы уж не мой верный раб как вы говорите ну здравствуйте **здравствуйте** я вижу что я вас пугаю садитесь и рассказывайте так говорила в июле года известная **анна павловна шерер** фрейлина и приближенная императрицы **мари** **феодоровны** встречая важного и чиновного князя **василия** первого приехавшего на ее вечер **анна павловна** кашляла несколько дней у нее был грипп как она говорила грипп был тогда новое слово употреблявшееся только редкими в записочках разосланных утром с красным лакеем было написано без различия во всех или если вас граф или князь нет в виду ничего лучшего и если перспектива вечера у бедной больной не слишком вас пугает то я буду очень рада видеть вас нынче у себя между семью и десятью часами **анна шерер** о какое жестокое нападение отвечал несколько не **смутясь** такою встречей вошедший князь в придворном шитом мундире в чулках башмаках при звездах с светлым выражением плоского лица он говорил на том изысканном французском языке на котором не только говорили но и думали наши деды и с теми тихими покровительственными интонациями которые свойственны состаревшемуся в свете и при дворе значительному человеку он подошел к **анне павловне** поцеловал ее руку подставив ей свою надушенную и сияющую лысину и покойно уселся на диване прежде всего скажите как ваше здоровье **успокойте** друга сказал он не изменяя голоса и тоном в котором **изза** приличия и участия просвечивало равнодушие и даже насмешка как можно быть здоровой когда нравственно страдаешь разве можно оставаться спокойною в наше время когда есть у человека чувство сказала **анна павловна** вы весь вечер у меня надеюсь а праздник английского посланника нынче среда мне надо показаться там сказал князь дочь заедет за мной и повезет меня я думала что нынешний праздник отменен признаюсь все эти праздники и фейерверки становятся несносны ежели бы знали что вы этого хотите праздник бы отменили сказал князь по привычке как заведенные часы говоря вещи которым он и не хотел чтобы верили не мучьте меня ну что же решили по случаю депеши **новосильцова** вы все знаете как вам сказать сказал князь холодным скучающим тоном что решили решили что **бонапарте** сжег свои корабли и мы тоже кажется готовы сжечь наши князь **василий** говорил всегда лениво как актер говорит роль старой пьесы **анна павловна шерер** напротив несмотря на свои сорок лет была преисполнена оживления и порывов быть энтузиасткой сделалось ее общественным положением и иногда когда ей даже того не хотелось она чтобы не обмануть ожиданий людей знавших ее делалась энтузиасткой сдержанная улыбка игравшая постоянно на лице **анны павловны** хотя и не шла к ее отжившим чертам выражала как у избалованных детей постоянное сознание своего милого недостатка от которого она не хочет не может и не находит нужным исправляться в середине разговора про политические действия **анна павловна** разгорячилась ах не говорите мне про австрию я ничего не понимаю может быть но австрия никогда не хотела и не хочет войны она предает нас россия одна должна быть спасительницей европы наш благодетель знает свое высокое призвание и будет верен ему вот одно во что я верю нашему **доброму** и чудному государю предстоит величайшая роль в мире и он так добродетелен и хорош что бог

3. Підготовка тексту

Проведено нормалізацію:

ё → е, ъ → ь;

перетворено у нижній регістр;

усі небуквені символи замінені на пробіл;

послідовності пробілів стиснуті до одного;

створено варіант тексту без пробілів.

4. Використані засоби

Для основних розрахунків використано Python-скрипт, який:

обчислює частоти символів та біграм;

знаходить H_1 , H_2 і надлишковість R ;

формує CSV-файли:

таблиці частот символів;

матриці частот біграм (для $\text{step}=1$, $\text{step}=2$);

для текстів зі пробілами і без.

5. Результати розрахунку H_1

Варіант тексту	H_1	Надлишковість R_1
З пробілами	4.38	0.122
Без пробілів	4.46	0.098

H_1 (з пробілом): 4.38595351 | $R_1 = 0.12280930$
 H_1 (без пробілів): 4.46859428 | $R_1 = 0.09801833$

З пробілами:

Символ	Кількість	Частота
	103318	0.160949013132273
о	61231	0.09538578972785194
а	45176	0.07037527456264701
е	42903	0.066834390042528
и	35789	0.05575218481765925
н	35097	0.054674185658872464
т	30590	0.04765317090648513
с	28102	0.04377735890206097
л	27261	0.04246724720763946
в	24794	0.03862414905051953
р	24545	0.0382362562896266
к	19315	0.030088950508622436
д	16380	0.025516800897294096
м	15921	0.02480176966335893
у	15443	0.02405714018662471
п	13838	0.021556867571230508
я	12469	0.019424236287445672
г	11168	0.017397535556836416
ь	10774	0.016783761469319085
ы	10223	0.015925412428146372
з	9592	0.014942439206767093
б	9302	0.014490676553518296
ч	7340	0.0114342685339523
й	6203	0.009663047372766501
ж	5456	0.008499369090087704
ш	5087	0.007924540058884925
х	4595	0.007158101350614553
ю	3494	0.005442961070521708
ц	2179	0.003394451108376303
э	1628	0.002536102067203589
щ	1511	0.0023538392036514884
ф	1206	0.001878709516613961

Без пробілів:

Символ	Кількість	Частота
о	61231	0.11368294802195272
а	45176	0.08387484868513884
е	42903	0.07965474218918256
и	35789	0.06644671860263046
н	35097	0.06516193475080392
т	30590	0.05679413009736137
с	28102	0.05217484942778847
л	27261	0.050613428590525276
в	24794	0.04603313702628237
р	24545	0.04557083763451241
к	19315	0.03586069378327999
д	16380	0.03041150215739716
м	15921	0.02955931171232724
у	15443	0.028671845410053992
п	13838	0.025691963788404267
я	12469	0.023150245445701174
г	11168	0.020734777539304732
ь	10774	0.020003267658351465
ы	10223	0.018980267799454897
з	9592	0.01780873801549167
б	9302	0.017270317037125053
ч	7340	0.013627620624865396
й	6203	0.011516639064855592
ж	5456	0.010129740889545722
ш	5087	0.009444646610175786
х	4595	0.00853118757101587
ю	3494	0.006487044477286061
ц	2179	0.004045583834002955
э	1628	0.0030225839751063847
щ	1511	0.002805358959696405
ф	1206	0.002239088620379791

6. Результати для біграм H_2

Методика:

Біграми формувалися у двох режимах:

step=1 — *перетинні біграми* (мають спільні символи);

step=2 — *неперетинні біграми* (беруться через два символи).

Варіант	step	H_2	R_2
З пробілами	1	3.97	0.204

Варіант	step	H_2	R_2
Без пробілів	1	4.15	0.162
З пробілами	2	3.97	0.204
Без пробілів	2	4.14	0.162

H2 step=1 (з пробілом): 3.97846804 | R2 = 0.20430639
H2 step=1 (без пробілів): 4.15065795 | R2 = 0.16219348
H2 step=2 (з пробілом): 3.97843012 | R2 = 0.20431398
H2 step=2 (без пробілів): 4.14979292 | R2 = 0.16236809

перетинні біграми 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
а	1.5578046	0.0006620	0.0034115	0.0009487	0.0021123	0.0008505	0.0011886	0.0039770	0.0001199	0.0007586	0.0042715	0.0088343	0.0026716	0.0050986	6.2312187	0.0007259	0.0026763	0.0037184	0.0046578	0.0001355	0.0005062	0.0009673	і
б	0.0011403	2.4924874	3.7387312	1.8693656	1.4020242	0.0019690	4.6734140	1.5578046	0.0007321	0.0	0.0002461	0.0007524	6.2312187	0.0002492	0.0022011	0.0	0.0011948	6.8543405	7.7890233	0.0011449	0.0	4.6734140	і
в	0.0057669	9.3468280	2.1809265	2.1809265	0.0002726	0.0041671	0.0	0.0005094	0.0032620	0.0	0.0001246	0.0009300	0.0001402	0.0013412	0.0069633	0.0003053	0.0004688	0.0029645	0.0002009	0.0007913	0.0	3.8945116	і
г	0.0010281	0.0	7.6332429	0.0	0.0008614	0.0005639	0.0	1.5578046	0.0008100	0.0	0.0001043	0.0016154	1.0904632	0.0002928	0.0089620	0.0	0.0013459	3.7387312	2.1809265	0.0007804	0.0	0.0	і
д	0.0040129	3.4271702	0.0009128	9.3468280	2.0251460	0.0042637	1.2462437	4.6734140	0.0027136	0.0	0.0002056	0.0006091	8.7237062	0.0016325	0.0034863	9.5026085	0.0020267	0.0002648	0.0002897	0.0015998	1.2462437	5.1407554	і
е	9.3468280	0.0010421	0.0013147	0.0033305	0.0024068	0.0014394	0.0009284	0.0010811	0.0001620	0.0022837	0.0009829	0.0059243	0.0037231	0.0066845	0.0002445	0.0009362	0.0067499	0.0034209	0.0035829	6.5427796	1.4020242	0.0010297	і
ж	0.0013350	4.9849749	0.0	2.3367070	0.0006231	0.0036966	1.2462437	0.0	0.0010920	0.0	8.4121452	1.5578046	1.0904632	0.0010203	3.4271702	0.0	0.0	2.0251460	0.0	0.0018840	0.0	0.0	і
з	0.0054118	0.0001433	0.0007461	0.0004299	0.0006620	0.0002071	0.0001713	7.7890233	0.0003115	1.5578046	0.0001620	0.0001760	0.0002554	0.0015079	0.0006854	3.4271702	0.0001448	0.0001464	1.2462437	0.0002570	0.0	0.0	і
и	7.0101210	0.0004829	0.0025843	0.0004735	0.0017120	0.0023289	0.0003847	0.0016434	0.0006231	0.0014129	0.0016512	0.0048852	0.0028087	0.0031732	0.0002695	0.0002165	0.0006620	0.0027775	0.0032854	3.1156093	1.7135851	0.0016668	і
й	1.5578046	1.4020242	0.0	3.1156093	0.0001791	1.2462437	0.0	0.0	0.0	1.5578046	4.0502921	0.0002243	4.6734140	0.0003396	1.2462437	0.0	5.6080968	0.0006496	0.0002990	0.0	0.0	0.0	і
к	0.0074883	1.5578046	0.0001666	3.1156093	1.8693656	0.0004735	1.0904632	7.7890233	0.0024940	0.0	9.3468280	0.0003551	1.5578046	0.0022588	0.0081878	0.0	0.0017104	7.0101210	0.0003941	0.0015141	0.0	3.1156093	і
л	0.0063947	2.3367070	9.3468280	0.0001479	0.0003271	0.0036468	0.0003302	6.2312187	0.0063059	0.0	0.0007804	0.0001152	9.3468280	0.0002492	0.0062265	0.0001028	0.0	0.0013443	4.2060726	0.0011714	7.7890233	0.0	і
м	0.0027432	3.1156093	0.0	6.0754382	1.5578046	0.0026903	0.0	3.1156093	0.0034318	0.0	6.6985601	7.4774624	1.5578046	0.0011169	0.0034972	0.0002227	4.2060726	0.0001308	5.6080968	0.0021700	1.5578046	1.5578046	і
н	0.0101927	1.0904632	9.3468280	0.0001651	0.0011185	0.0086317	2.6482679	1.7135851	0.0065458	0.0	0.0002180	3.1156093	3.1156093	0.0028492	0.0083124	1.2462437	3.8945116	0.0003987	0.0005935	0.0026311	9.3468280	0.0	і
о	7.7890233	0.0030377	0.0079899	0.0046204	0.0044303	0.0020952	0.0015858	0.0008723	0.0007757	0.0035720	0.0016652	0.0067156	0.0048167	0.0066689	0.0001853	0.0010935	0.0061502	0.0064680	0.0066642	5.2965359	0.0004642	0.0005390	і
п	0.0011527	0.0	1.4020242	0.0	0.0024192	0.0	0.0007446	0.0	0.0007446	0.0	6.0754382	0.0007088	0.0	5.1407554	0.0084915	7.9448038	0.0059118	3.1156093	4.8291945	0.0006792	1.5578046	0.0	і
р	0.0082999	0.0001215	0.0002554	0.0002772	0.0003972	0.0059196	0.0003645	2.713898	0.0051781	0.0	0.0002523	6.0754382	0.0002305	0.0006418	0.0068216	4.3618531	4.9849749	0.0002352	0.0006371	0.0030423	2.1809265	0.0001308	і
с	0.0013989	4.2060726	0.0017042	1.0904632	0.0003053	0.0033274	2.1809265	4.9849749	0.0016434	0.0	0.0046438	0.00022915	0.0009689	0.0007025	0.0024410	0.0016746	0.0002585	0.0007430	0.0103749	0.0007165	1.7135851	0.0001822	і
т	0.0053557	2.8040484	0.0020999	9.3468280	0.0001292	0.0045378	1.5578046	6.2312187	0.0033383	0.0	0.0003442	0.0002695	4.6734140	0.0008723	0.0132148	6.0754382	0.0027947	0.0008053	0.0001059	0.0017447	1.4020242	2.1809265	і
у	1.8693656	0.0005047	0.0003818	0.0012244	0.0019098	0.0001464	0.0012509	0.0008583	1.7135851	9.0352671	0.0009876	0.0014845	0.0009798	0.0003022	1.2462437	0.0005452	0.0005670	0.0013848	0.0014207	4.6734140	6.2312187	0.0003551	і
ф	0.0001775	0.0	0.0	0.0	0.0001370	0.0	0.0006277	0.0	0.0006277	0.0	1.0904632	9.1910476	4.6734140	0.0	5.7638773	0.0	0.0004676	6.2312187	7.7890233	7.0101210	4.6734140	0.0	і
х	0.0009658	0.0	0.0001277	0.0	0.0	3.8945116	0.0	0.0	0.0001526	0.0	0.0	6.0754382	5.232163	0.0001479	0.0022416	1.5578046	7.3216819	8.945116	2.4924874	0.0001199	0.0	0.0	і
ц	0.0005623	0.0	1.8693656	2.6482679	3.1156093	0.0107170	0.0	0.0	0.0002227	0.0	0.0001713	4.6734140	0.0	2.6482679	0.0004097	0.0	1.5578046	3.1156093	0.0004860	0.0	0.0	0.0	і
ш	0.0021123	0.0	0.0	0.0	0.0	0.0029380	0.0	0.0	0.0011621	0.0	0.0002227	3.5829507	0.0	0.0004564	9.3468280	0.0	2.0251460	0.0	0.003991	0.0005701	0.0	0.0	і
щ	0.0013646	0.0	1.5578046	0.0	0.0	0.0021560	0.0	0.0	0.0019098	0.0	0.0004081	0.0004548	2.0251460	0.0003286	0.0003271	5.4523163	3.1156093	1.5578046	0.0001588	0.0002850	0.0	0.0	і

перетинні біграми 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
а	0.0003861	0.0015837	0.0058205	0.0017192	0.0032676	0.0015317	0.0015335	0.0051911	0.0015707	0.0009060	0.0065167	0.0108705	0.0040938	0.0788880	0.0014723	0.0027125	0.0037336	0.0063292	0.0064053	0.0005551	0.0007129	0.0012773	0
б	0.0013683	3.5275922	9.9412080	2.2279530	2.0422902	0.0023597	9.2831375	3.7132550	0.0009041	0.0	0.0003026	0.0008967	8.1691610	0.003156	0.0027032	5.5698825	0.0014258	9.6544630	1.1139765	0.0013831	0.0	5.7555452	7
в	0.0069864	0.0002487	0.0004641	0.0004121	0.0007073	0.0051261	2.9706040	0.0007593	0.0041978	0.0	0.0007333	0.0012959	0.0003490	0.0021146	0.0087632	0.0010619	0.0008020	0.0042479	0.0007946	0.0010879	1.6709647	7.9834982	7
г	0.0012402	4.2702432	0.0001652	1.2996392	0.0010991	0.0006980	0.0	7.6121727	0.0010304	0.0	0.0001875	0.0019271	5.0128942	0.0004437	0.0107665	9.8401258	0.0016282	0.0001559	4.8272315	0.0009617	7.4265100	3.7132550	и
д	0.0047993	9.4688003	0.0011752	9.2831375	5.5698825	0.0051112	2.2279530	5.1985570	0.0033289	0.0	0.0003360	0.0007426	0.0001485	0.0020719	0.0042182	0.0002376	0.0024581	0.0004604	0.0003880	0.0019383	1.9412995	6.4981962	0
е	0.0001949	0.0023890	0.0034421	0.0045450	0.0037243	0.0020478	0.0013367	0.0019643	0.0014481	0.0027218	0.0021833	0.0074562	0.0053546	0.0094428	0.0016115	0.0031395	0.0085404	0.0005053	0.0005829	6.8695217	0.0014277	0	
ж	0.0015911	6.6838590	2.9706040	3.3419295	0.0007556	0.0044187	1.4853020	9.2831375	0.0013144	0.0	0.0001132	1.8566275	2.9706040	0.0012439	5.7555452	1.4853020	0.0	3.1562667	1.2996392	0.0002830	0.0	5.5698825	0
з	0.0064610	0.0002005	0.0009765	0.0005625	0.0008689	0.0002580	0.0002079	0.0001151	0.0042233	1.8566275	0.0003379	0.00020413	0.0003471	0.0019067	0.0009041	0.0001912	0.0002376	0.0003416	7.2840872	0.0003304	3.7132550	5.5698825	1
и	0.003880	0.0014240	0.0052486	0.0012810	0.0029798	0.0032695	0.0006126	0.0025454	0.0019921	0.0016895	0.0031933	0.0061268	0.0039360	0.0057017	0.0002701	0.0025621	0.0014351	0.0053526	0.0047139	0.0005551	0.0001448	0.0001388	0
й	0.0020005	0.0004420	0.0007389	0.0003731	0.0007185	0.0001262	0.0001225	0.0001615	0.0006015	1.8566275	0.0008020	0.0004140	0.0004604	0.0009654	0.0006256	0.0008484	0.0004158	0.0017359	0.0006498	0.0002469	5.9412080	5.0128942	5
к	0.0090213	0.0006442	0.0006219	0.0002097	0.0002989	0.0006479	0.0001670	9.4688003	0.0032973	1.8566275	0.0004493	0.0004945	0.0002302	0.0032156	0.0101910	0.0003936	0.0021351	0.0006238	0.0007240	0.0019142	5.5992785	0.0128942	3
л	0.0077477	0.0004158	0.0008076	0.0004660	0.0008076	0.0048179	0.0004307	0.0001912	0.0079575	1.8566275	0.0018584	0.0002636	0.0001281	0.011046	0.0085386	0.0009246	0.0003583	0.0023727	0.0003694	0.0015443	5.5275922	2.9706040	5
м	0.0034943	0.0002952	0.0008169	0.0004455	0.0003601	0.0034943	0.0001206	0.0001986	0.0047993	0.0	0.0007092	0.0002914	0.0002729	0.0020292	0.0047789	0.010304	0.0002525	0.0010731	0.0003404	0.0029425	1.985570	9.2831375	4
н	0.0122203	0.0003434	0.0005532	0.0003430	0.0015335	0.003581	9.412080	0.0002116	0.0081765	0.0	0.0005196	3.899177	0.001243	0.0039676	0.0102875	0.0006813	0.0001670	0.0009988	0.0009227	0.003270	2.4136157	5.5698825	0
о	0.0002302	0.0005426	0.0120346	0.0005605	0.0062809	0.0031335	0.0023040	0.0016561	0.002947	0.0042572	0.0032231	0.0083696	0.0065631	0.0104305	0.0023449	0.0035424	0.0078959	0.0009008	0.000287	0.0005970	0.0006925	0.0007946	0
п	0.0013731	1.8566275	1.8566275	1.6709647	1.5566275	0.0028830	0.0	0.0008911	0.0	0.0	7.9834982	0.0008447	0.0	6.3125305	0.012609	4.688003	0.0070459	1.1399765	1.1268707	0.0009480	1.8566275	0.0	1
р	0.0093885	0.0002079	0.0004363	0.0003676	0.0051918	0.0070756	0.0004381	6.6838590	0.0028464	1.8566275	0.0006486	9.118120	0.0003063	0.0009688	0.0083240	0.0004239	2.813375	0.0004039	0.0003691	0.0008391	4.8723151	0.0001745	0
с	0.0017359	0.0002135	0.0002743	0.0001336	0.0001254	0.0004304	0.0001058	0.0002413	0.0020627	1.8566275	0.0005741	0.0028090	0.0012476	0.0012272	0.0030875	0.0023077	0.0004455	0.0012717	0.0126900	0.0001304	1.985570	0.0002302	2
т	0.0064926	0.0002211	0.0002928	0.0001596	0.0004066	0.0005930	0.0018612	0.0001392	0.0043296	3.7132550	0.0005475	0.0003954	0.0002432	0.0018502	0.0106691	0.0005458	0.0004310	0.0004500	0.0003787	0.0022224	3.6415687	9.412080	3
у	0.0001968	0.0003891	0.0015985	0.0016951	0.0025992	0.0002822	0.0014514	0.001752	0.0006683	0.0001113	0.0008540	0.018677	0.0015224	0.0011008	0.0004474	0.0013126	0.0008420	0.0023540	0.0020181	0.0001764	8.7273151	0.0004697	1
ф	0.0002172	3.7132550	2.5997885	5.5698825	5.5698825	0.0001652	0.0	7.4265100	0.0007723	0.0	2.7894412	0.001151	1.1139765	2.4136157	7.793835	2.2279530	5.9706042	1.6709647	2.813375	5.5698825	0.0	0	
х	0.0011938	0.0001762	0.0004845	0.0001726	0.0001502	0.0001584	2.704232	7.9834982	0.0004827	1.8566275	0.0002506	0.0002172	0.0002079	0.0004063	0.0029316	0.0003917	0.0002190	0.0004381	0.0001448	0.0002209	4.161897	2.813375	0
ц	0.0006739	2.813375	5.1985570	3.899177	2.0422902	0.0012884	1.8566275	9.2831375	0.0002877	0.0	0.0002190	2.813375	2.2279530	5.5698825	0.0005161	1.899877	5.5698825	0.0048505	1.8566275	0.0005829	3.7132550	0	
ш	0.0025250	1.4853020	1.1139765	5.5698825	5.5698825	0.0003015	0.0	5.5698825	0.0014017	0.0	0.0002766	4.2704235	5.5698825	0.0005005	0.0001336	3.2996392	3.3419295	2.7894412	0.0004604	0.0006790	1.8566275	0	
щ	0.0016301	1.6709647	3.1562667	0.0	1.2996392	0.0025995	0.0	1.8566275	0.002836	0.0	0.0004901	0.0005423	3.4191295	0.0004084	0.0003954	8.1691610	5.5698825	1.2996392	0.0001912	0.0003397	1.8566275	0	
ч	0.0004827	0.0	0.0	0.0	1.8566275	0.0014036	0.0	0.0	0.0007723	0.0	0.0	0.0	0.0	0.0	3.5275922	0.0	3.7132550	1.8566275	5.5698825	7.793835	0.0	1	

	а	б	в	г	д	е	ж	з	и	й	к	л	м	н	о	п	р	с	т	у	ф	х	ц
а	1.8693627	0.0007321	0.0034988	0.0008848	0.0020781	0.0008661	0.0012182	0.0039381	0.0001308	0.0007415	0.0042122	0.0008845	0.0027168	0.0050503	6.2312090	0.0006916	0.0026763	0.0037605	0.0046484	0.0001557	0.0004704	0.0009596	6.
б	0.0011808	2.1809231	2.8040440	6.2312090	1.8693627	0.0019780	6.2312090	3.1156045	0.0007165	0.0	0.0002212	0.0007726	6.2312090	0.0002554	0.0021778	0.0	0.0011060	7.1658903	3.468135	0.0011963	0.0	4.3618463	6.
в	0.0056112	3.1156045	2.4924836	2.1809231	0.0002804	0.0041811	0.0	0.0005234	0.0032589	0.0	0.0001402	0.0009409	0.0001246	0.0012836	0.0070474	0.0003240	0.0004579	0.0030439	0.0001931	0.0008256	0.0	4.0502858	9.
г	0.0010343	0.0	6.5427694	0.0	0.0008318	0.0006324	0.0	0.0	0.0008349	0.0	0.0001246	0.0016481	1.2462418	0.0003520	0.00089604	0.0	0.0013083	3.7387254	3.1156045	0.0007446	0.0	0.0	0.
д	0.0040907	3.1156045	0.0009222	6.2312090	1.5578022	0.0043026	1.8693627	3.1156045	0.0006420	0.0	0.0002399	0.0005888	8.1005717	0.0017229	0.0034583	9.9693444	0.0020313	0.0002461	0.0003146	0.0015731	1.2462418	5.2965276	0.
е	3.1156045	0.0010250	0.0012649	0.0033087	0.0023553	0.0014736	0.0009471	0.0010374	0.001775	0.0023367	0.0009720	0.0059570	0.0036390	0.0067795	0.0002274	0.0009720	0.0066113	0.0034489	0.0034427	6.5427694	1.8693627	0.0010437	0.
ж	0.0012431	4.6734067	0.0	3.4271649	0.0006418	0.0036234	1.2462418	0.0	0.0011309	0.0	9.3468135	0.0	6.2312090	0.010312	1.8693627	0.0	0.0	1.8693627	0.0	0.001775	0.0	0.0	0.
з	0.0053339	0.0001308	0.0007197	6.0004361	0.0006792	0.0002399	0.0001588	8.7236926	0.0003084	0.0	0.0001588	0.0001588	0.0002367	0.0015266	0.0007290	4.0502858	0.0001277	0.0001536	1.2462418	0.0002149	0.0	0.0	3.
и	5.2965276	0.0005296	0.0025984	0.0004860	0.0017135	0.0023180	0.0003832	0.0016076	0.0005732	0.0014020	0.0016450	0.0047544	0.0028757	0.0031498	0.0002585	0.0002212	0.0006728	0.0002809	0.0032682	3.1156045	2.1809231	0.0016948	0.
й	0.0	1.2462418	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	4.0502858	0.0002274	3.4271649	0.0003240	1.2462418	0.0	5.2965276	0.0006605	0.0003115	0.0	0.0	0.0	3.
к	0.0076706	0.0	0.0001620	3.1156045	1.8693627	0.0004611	1.2462418	1.2462418	0.0023304	0.0	6.2312090	0.0003582	0.0	0.0023491	0.0081130	0.0	0.0016699	7.4774508	0.0004143	0.0015609	0.0	3.1156045	6.
л	0.0064181	1.5578022	1.5578022	0.0001339	0.0003302	0.0036483	0.0003614	6.2312090	0.0063994	0.0	0.0007290	0.0001246	9.3468135	0.0002336	0.0061657	0.0001152	0.0	0.0013864	4.3618463	0.0012119	3.468135	0.0	0.
м	0.0026420	6.2312090	0.0	5.9196485	0.0	0.0026482	0.0	3.1156045	0.0033492	0.0	6.5427694	9.6583739	1.2462418	0.012213	0.0035954	0.0002243	4.3618463	0.0004393	0.0005389	0.0026731	6.2312090	0.0	0.
н	0.0103531	9.3468135	1.8693627	0.0001775	0.0010935	0.0086551	3.4271649	1.5578022	0.0064274	0.0	0.0002554	0.0	3.1156045	0.0028383	0.0082781	6.2312090	4.0502858	0.0004393	0.0005389	0.0026731	6.2312090	0.0	0.
о	6.2312090	0.0031623	0.0077959	0.0047045	0.0044895	0.0021341	0.0015702	0.0009440	0.0007228	0.0034832	0.0016730	0.00065427	0.0049039	0.0065645	0.0001526	0.0010904	0.0061470	0.0064866	0.0006892	5.2965276	0.0004424	0.0005795	9.
п	0.0011029	0.0	9.3468135	0.0	0.0024644	0.0	0.0	0.0007446	0.0	0.0	7.4774508	0.0006978	0.0	4.9849672	0.0085679	6.5427694	0.0059040	6.2312090	7.1658903	0.0006854	3.1156045	0.0	3.
р	0.0083529	0.0001526	0.0002679	0.0004206	0.0059508	0.0004019	4.9849672	0.0053077	0.0	0.0002990	4.3618463	0.0001775	0.0006324	0.0068013	4.0502858	0.0002274	0.0006231	0.0002740	0.0006231	1.8693627	0.0001246	0.0	0.
с	0.0014176	4.9849672	0.0016948	6.2312090	0.0002928	0.0033212	1.2462418	7.1658903	0.0017011	0.0	0.0045020	0.0023304	0.0009471	0.0007072	0.0024052	0.0016730	0.0002554	0.0007352	0.0102846	0.0007321	1.5578022	0.0002056	1.
т	0.0053806	2.8040440	0.0021123	3.1156045	0.0001090	0.0044521	0.0	6.2312090	0.0032807	0.0	0.0003520	0.0002617	5.6080881	0.0008412	0.0132195	6.8543299	0.0028850	0.0007352	0.0001090	0.0017166	2.4924836	2.4924836	0.
у	2.1809231	0.0005670	0.0008100	0.0012306	0.0018942	0.0001183	0.0012213	0.0008194	1.5578022	0.0001059	0.0009907	0.0015110	0.010374	0.0030353	2.1809231	0.0005701	0.0005795	0.0013615	0.0013864	3.1156045	6.2312090	0.0003614	0.
ф	0.0001838	0.0	0.0	0.0	0.0001277	0.0	0.0	0.0006698	0.0	1.8693627	7.7890112	0.0	0.0	5.2965276	0.0	0.0004735	9.3468135	6.2312090	8.4121321	3.1156045	0.0	0.0	0.
х	0.0009720	0.0	0.0001090	0.0	4.0502858	0.0	0.0	0.0001588	0.0	0.0	6.5427694	5.9196485	0.0001620	0.0021778	0.0	7.4774508	3.7387254	2.1809231	0.0001246	0.0	0.0	0.0	0.
ц	0.0005888	0.0	1.8693627	2.1809231	3.1156045	0.0011029	0.0	0.0002617	0.0	0.0001651	3.1156045	0.0	3.7387254	0.0004081	0.0	0.0	3.1156045	0.0005047	0.0	0.0	0.0	0.0	0.
ч	0.0021154	0.0	0.0	0.0	0.0029878	0.0	0.0	0.0011060	0.0	0.0002087	4.6734067	0.0	0.0004455	9.6583739	0.0	1.8693627	0.0	0.0035455	0.0005950	0.0	0.0	0.0	0.
ш	0.0014331	0.0	1.2462418	0.0	0.0021279	0.0	0.0	0.0019441	0.0	0.0004393	0.0004112	2.4924836	0.0003738	0.0003364	6.8543299	6.2312090	0.0	0.0002180	0.0002772	0.0	0.0	0.0	0.
щ	0.0003925	0.0	0.0	0.0	0.0012836	0.0	0.0	0.0007072	0.0	0.0	0.0	0.0	0.0	2.4924836	0.0	3.1156045	0.0	0.0	0.0	7.7890112	0.0	0.0	0.
ы	0.0	0.0004860	0.0010935	0.0001059	8.4121321	0.0009253	3.1156045	5.2965276	1.2462418	0.0014456	6.0001744	0.0002962	0.0014799	0.0003240	6.2312090	0.0001308	0.0004081	0.0008194	0.0005265	3.1156045	0.0	0.0008287	0.
я	1.8693627	0.0007321	0.0034988	0.0008848	0.0020781	0.0008661	0.0012182	0.0039381	0.0001308	0.0007415	0.0042122	0.0008845	0.0027168	0.0050503	6.2312090	0.0006916	0.0026763	0.0037605	0.0046484	0.0001557	0.0004704	0.0009596	6.

неперетинні біграми 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
а	0.0004530	0.0015892	0.0057109	0.016672	0.0031971	0.0015521	0.0015447	0.0052319	0.0016152	0.0009245	0.0065278	0.0108761	0.0040363	0.0080985	0.0013479	0.0026104	0.0036612	0.0063905	0.0062828	0.0005829	0.0006906	0.0012699	6.
б	0.0013961	2.9705985	7.0551714	1.8566240	7.4264962	0.0024581	7.4264962	3.7132481	0.0009171	0.0	0.0003156	0.0008540	8.1691458	0.0003230	0.0027478	1.1139744	0.0015521	0.0001002	1.8566240	0.0014593	0.0	5.1985473	3.
в	0.0070625	0.0002487	0.0004790	0.0004270	0.0006720	0.0050203	2.5992736	0.0008057	0.0040950	0.0	0.0007092	0.0012884	0.0003601	0.0021759	0.0087929	0.010248	0.0007834	0.0042331	0.0008354	0.0010842	1.8652992	7.0551714	7.
г	0.0012847	4.0845729	0.0001559	1.1139744	0.0011325	0.0006758	0.0	7.0551714	0.0010434	0.0	0.0002320	0.0018863	4.8272225	0.0004121	0.0107721	0.0001188	0.0015669	0.0001485	6.3125218	0.0009394	7.4264962	3.7132481	0.
д	0.0047640	0.0001002	0.0011808	0.0001039	7.4264962	0.0051688	2.2279488	5.5698721	0.0033679	0.0	0.0003527	0.0007573	0.0001411	0.0021462	0.0043259	0.0002265	0.0024841	0.0004975	0.0003898	0.0019643	9.7905985	5.9411970	0.
е	0.0001819	0.0022242	0.0034533	0.0046341	0.0035610	0.0020311	0.0012736	0.0020608	0.0014704	0.0026809	0.0002276	0.0075004	0.0056552	0.0093722	0.0016523	0.0031896	0.0084402	0.0061120	0.0050945	0.0005755	9.411970	0.0014407	0.
ж	0.0017155	6.6838462	2.2279488	2.2279488	0.0007760	0.0043816	1.8566240	7.4264962	0.0012959	0.0	0.0001002	0.0	2.2279488	0.0012773	5.1985473	7.4264962	0.0	2.5992736	1.8566240	0.0002636	0.0	7.4264962	7.
з	0.0065427	0.0001745	0.0009431	0.0005681	0.0008280	0.0002599	0.0001708	0.0001039	0.0004418	0.0	0.0003490	0.0001856	0.0003156	0.0018751	0.0009394	0.0002042	0.0002302	0.0004047	5.1985473	0.0002970	0.0	3.7132481	7.
и	0.0004233	0.0014890	0.0053136	0.0012476	0.0030225	0.0031302	0.0005569	0.0025472	0.0021091	0.0017749	0.0032416	0.0006934	0.0038729	0.0057555	0.0019234	0.0025732	0.0013553	0.0052839	0.0047863	0.0005718	0.0001336	0.0021722	0.
й	0.0002116	0.0003973	0.0008057	0.0003119	0.0007055	0.0001411	0.0001002	0.0001151	0.0006201	3.7132481	0.0008689	0.0003898	0.0004790	0.0009877	0.0006163	0.0008726	0.0003638	0.0017378	0.0006572	0.0002636	5.9411970	4.4558977	0.
к	0.0087744	0.0006126	0.0006795	0.0002339	0.0003230	0.0006052	0.0001633	8.9117955	0.0034421	3.7132481	0.0004233	0.0005050	0.0002005	0.0032453	0.0101445	0.0003490	0.0020237	0.0006201	0.0007315	0.0018751	2.9705985	5.5698721	4.
л	0.0077941	0.0003936	0.0008466	0.0004641	0.0008429	0.0004860	0.0004418	0.0001930	0.0079797	3.7132481	0.0019457	0.0002562	0.0001076	0.011251	0.0004592	0.0009543	0.0004084	0.0002021	0.0003564	0.0015595	3.7132481	2.5992736	7.
м	0.0033976	0.002970	0.0008354	0.0004418	0.0003638	0.0033270	0.001336	0.0022777	0.0004759	0.0	0.0005890	0.0002784	0.0002673	0.0019680	0.001497	0.0010174	0.0002562	0.0016619	0.0003861	0.0027812	4.8272225	0.0001002	0.
н	0.0124245	0.0011939	0.0004975	0.003267	0.0051705	0.010888	3.125218	0.0002153	0.0008037	0.0	0.0006050	0.0845729	0.0001559	0.0005085	0.0105084	0.0006423	0.0001708	0.0009095	0.0008948	0.0002342	5.992736	5.1985473	0.
о	0.0002116	0.0005203	0.0119975	0.0006407	0.0006079	0.0031674	0.023579	0.0016486	0.0023133	0.0041811	0.0033939	0.0008198	0.006518	0.0105604	0.0022910	0.0034904	0.0079941	0.0092787	0.001271	0.0006052	0.0006015	0.0007240	0.
п	0.0013404	3.7132481	0.0	1.1139744	3.7132481	0.0029233	0.0	0.0009208	0.0	8.5404706	0.0009849	0.0	5.9411970	0.0102337	8.5404706	0.0069317	1.1139744	7.0551714	0.0008726	0.0	0.0	0.0	0.
р	0.0098066	0.0001819	0.0004641	0.0004195	0.0005012	0.0073039	0.0004307	5.5698721	0.0062642	3.7132481	0.0004084	0.001706	0.0002970	0.0008516	0.0087112	0.0003044	6.6838462	0.0004530	0.0007946	0.0036612	6.3125218	0.0001596	0.
с	0.0017415	0.0002059	0.0023216	0.0001633	0.0005005	0.0039991	0.0001076	0.0002025	0.0027019	0.0	0.0057926	0.0026809	0.0013256	0.0012030	0.0030522	0.0024136	0.0003780	0.0017173	0.0126176	0.0014824	4.558977	0.0002143	0.
т	0.0064870	0.0003067	0.0002954	0.0001782	0.0004195	0.0056292	6.6838462	0.0001262	0.0042516	3.7132481	0.0007166	0.0004010	0.0002487	0.0041258	0.0160895	0.0005161	0.0034356	0.010159	0.0004084	0.0021725	9.411970	5.5698721	9.
у	0.001893	0.0004606	0.0016449	0.0017749	0.0025324	0.0002896	0.0051075	0.001251	0.0025904	0.0001299	0.0018009	0.0019011	0.0014815	0.001585	0.0004451	0.002402	0.0007575	0.0023890	0.0020942	0.0001596	4.0845729	0.0005012	0.
ф	0.0002190	3.7132481	2.5992736	0.0	7.4264962	0.0001596	0.0	7.4264962	0.0002060	0.0	3.1492333	0.0012193	1.4852992	0.0001298	6.6838462	2.2279488	0.0003032	2.9705985	1.4852992	0.0001076	1.1139744	0.0	0.
х	0.0012030	0.0001306	0.0004084	0.001745	0.0001411	0.0001131	4.8272225	0.0001133	0.0005272	3.7132481	0.0001930	0.0002199	0.0002376	0.0004344	0.0030003	0.0003638	0.0002032	0.0004121	0.0002190	1.2279488	3.7132481	0.	0.
ц	0.0007203	1.1139744	4.8272225	0.0845729	2.2279488	0.0034739	3.7132481	3.7132481	0.0002673	0.0	0.0002421	4.852992	2.9705985	5.1985473	0.0005590	2.9705985	3.7132481	4.0845729	2.2279488	0.0005866	0.0	3.7132481	0.
ч	0.0025955	1.1139744	3.7132481	3.7132481	0.0001336	0.0004100	7.4264962	0.0013553	0.0	0.0002784	3.4192333	2.7496962	0.0005012	0.0010559	1.4852992	2.5992736	2.5992736	0.0004077	0.0006683	0.0	0.0	0.0	0.
ш	0.0015707	1.1139744	3.4192333	0.0	1.4852992	0.0024804	0.0	0.00022465	0.0	0.0004752	0.0005384	4.4558977	0.0004270	0.0004455	8.1691458	3.7132481	7.4264962	0.0001968	0.0003710	0.0	0.0	0.0	0.
щ	0.00048270	0.0	0.0	0.0	3.7132481	0.0037760	0.0	0.00023890	0.0	0.0	0.0	0.0	0.0	0.0	3.4192333	0.0	0.0	3.7132481	0.0	7.0551714	0.0	0.0	0.

При переході до $\text{step}=2$ (неперетинні біграми) значення H_2 дещо збільшується.

При збільшенні порядку n у CoolPinkProgram $H^{(n)}$ зменшується (програма “вчиться” краще передбачати наступний символ).

Вилучення пробілів робить розподіл рівномірнішим $\rightarrow H$ трохи зростає, надлишковість зменшується.

9. Висновки

1. Проведено розрахунок ентропій H_1 , H_2 та $H^{(n)}$ на тексті російською мовою.
2. Отримані результати узгоджуються з теорією Шеннона: зі збільшенням порядку n умовна ентропія прямує до деякої межі ($\sim 1\text{--}2$ біт/символ).
3. Надлишковість природної мови становить близько 40–60 %, що свідчить про значну структурованість текстів.
4. Алгоритм у Python та експеримент у CoolPinkProgram підтвердили взаємну залежність символів у природних мовах.