

НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ УКРАЇНИ
«КИЇВСЬКИЙ ПОЛІТЕХНІЧНИЙ ІНСТИТУТ ІМЕНІ ІГОРЯ
СІКОРСЬКОГО»

Навчально-науковий фізико-технічний інститут
КАФЕДРА ІНФОРМАЦІЙНОЇ БЕЗПЕКИ

«Криптографія»

Комп'ютерний практикум №1

Студенти: Маврикін Едуард

Слобода Ірина

Група: ФБ-25

Варіант 2

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.

При написанні коду особливих проблем не виникло, єдине - вивести таблицю частот біграм повністю неможливо через її розмір, тому вона додатково зберігається у CSV файл. Також після виконання основного коду можна вводити букви, щоб отримати найімовірнішу наступну літеру; для цього використовуються частоти біграм, які перетинаються, та алфавіт із пробілом.

Таблиця з отриманими значеннями ентропії:

	З пробілами	Без пробілів
H_1 (ентропія одиночних букв)	4.340	4.451
H_2 (біграми без перекриття)	3.970	4.127
H_2 (біграми з перекриттям)	3.970	4.127

Таблиці з оцінками надлишковості російської мови (R):

	З пробілами	Без пробілів
R_1 (на основі H_1)	13.95%	10.99%
R_2 (біграми без перекриття)	21.29%	17.45%
R_2 (біграми з перекриттям)	21.30%	17.45%

З пробілами ($H_0 = 5.044$):

$$R_1 = 1 - 4.3401/5.044 = 0.1395 = 13.95\%$$

$$R_2 \text{ (без перекриття)} = 1 - 3.9701/5.044 = 0.2129 = 21.29\%$$

$$R_2 \text{ (з перекриттям)} = 1 - 3.9697/5.044 = 0.2130 = 21.30\%$$

Без пробілів ($H_0 = 5.000$):

$$R_1 = 1 - 4.4507/5.000 = 0.1099 = 10.99\%$$

$$R_2 \text{ (без перекриття)} = 1 - 4.1275/5.000 = 0.1745 = 17.45\%$$

$$R_2 \text{ (з перекриттям)} = 1 - 4.1274/5.000 = 0.1745 = 17.45\%$$

Приклад роботи програми

```
Sorted letter frequency with spaces:
(letter, count, frequency)
```

```
179144 0.1753391406479397
о 96681 0.09462758148184398
е 73380 0.0718214740138984
а 67115 0.06568953704609964
н 54835 0.053670353332680824
и 54644 0.053483410002936285
т 54553 0.05339434276206323
с 44600 0.04365273563668396
в 38984 0.03815601448566115
л 38734 0.037911324263482434
р 35246 0.03449740628364491
к 27833 0.02724185181560145
д 26983 0.026409905060193795
м 26495 0.02593226974650093
у 24993 0.024462170891651168
п 23121 0.0226299305079769
ь 19362 0.01895076832729764
я 18000 0.017617695996867966
ч 15249 0.01492512479201331
б 14655 0.014343740824116668
г 14232 0.013929724968190272
ы 13912 0.013616521483801507
э 12975 0.012699422531075658
ж 9615 0.009410785944993639
й 8437 0.008257805618087501
х 7171 0.007018694332974454
ш 6937 0.006789664285015171
ю 4733 0.00463247528628756
э 2971 0.002907898600371929
щ 2521 0.00246745620045023
ц 2336 0.002286385436037976
ф 1049 0.0010267201722619164
ъ 204 0.00019966722129783693
```

H1 (entropy of single letters): 4.340109159286999
H2 (entropy of bigrams without overlapping): 3.9700541960289835
H2 (entropy of bigrams with overlapping): 3.9697042983305573

Bigram frequency matrix without overlapping:																					
	а	б	в	г	д	е	...	ч	ш	щ	ь	ы	б	э	ю	я					
а	0.021742	0.002435	0.006370	0.015537	0.002688	0.007278	0.003745	0.002044	0.003978	...	0.005724	0.000499	0.000051	0.000002	0.000000	0.002664	0.000045	0.002582	...		
б	0.015572	0.000037	0.000550	0.002649	0.000685	0.001779	0.001149	0.001609	0.003414	...	0.000953	0.000859	0.000260	0.000000	0.000000	0.000014	0.000073	0.001959	...		
в	0.000413	0.000640	0.000002	0.000045	0.000000	0.000020	0.002136	0.000006	0.000002	...	0.000018	0.000010	0.000143	0.000149	0.003602	0.000057	0.000012	0.000008	0.000017	...	
г	0.005168	0.005540	0.000025	0.000059	0.000045	0.000844	0.005090	0.000008	0.000507	...	0.000192	0.000574	0.000000	0.000004	0.002703	0.000147	0.000020	0.000000	0.000166	...	
д	0.000912	0.000959	0.000002	0.000020	0.000000	0.001088	0.000190	0.000002	0.000004	...	0.000025	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	...	
е	0.000787	0.005129	0.000035	0.000910	0.000004	0.000039	0.004269	0.000022	0.000016	...	0.000053	0.000069	0.000000	0.000000	0.000000	0.000427	0.001092	0.000000	0.000014	0.000431	...
...	0.016915	0.000016	0.001298	0.001247	0.003351	0.002672	0.001846	0.000799	0.001221	...	0.001186	0.000969	0.000949	0.000000	0.000000	0.000000	0.000016	0.000211	0.000217	0.000000	...
ч	0.000671	0.001198	0.000020	0.000006	0.000002	0.000673	0.004436	0.000010	0.000000	...	0.000008	0.000000	0.000000	0.000000	0.000000	0.000047	0.000002	0.000000	0.000000	0.000000	...
ш	0.001071	0.004444	0.000123	0.000005	0.000139	0.000738	0.000164	0.000051	0.000041	...	0.000031	0.000006	0.000000	0.000041	0.000227	0.000114	0.000000	0.000000	0.000000	0.000442	...
щ	0.015864	0.000078	0.000668	0.002584	0.000548	0.001734	0.001683	0.000215	0.001682	...	0.001482	0.000015	0.000135	0.000000	0.000000	0.000000	0.000020	0.000301	0.001061	0.001861	...
ь	0.005458	0.000004	0.000033	0.000037	0.000031	0.000268	0.000000	0.000010	0.000010	...	0.000206	0.000178	0.000002	0.000000	0.000000	0.000000	0.000006	0.000000	0.000000	0.000004	...
ы	0.004406	0.006734	0.000027	0.000258	0.000096	0.000012	0.000362	0.000010	0.000014	...	0.000012	0.000014	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	...
б	0.006503	0.005311	0.000031	0.000049	0.000145	0.000033	0.003565	0.000319	0.000018	...	0.000254	0.000000	0.000000	0.000000	0.000000	0.000009	0.004659	0.000002	0.000681	0.001157	...
э	0.006763	0.002834	0.000043	0.000045	0.000096	0.000043	0.003923	0.000006	0.000025	...	0.000051	0.000008	0.000000	0.000000	0.000000	0.000078	0.000006	0.000006	0.000000	0.000160	...
ю	0.003999	0.009991	0.000018	0.000041	0.000031	0.000298	0.009976	0.000002	0.000025	...	0.000186	0.000006	0.000098	0.000000	0.000000	0.002463	0.000006	0.000000	0.000000	0.001772	...
я	0.021676	0.000012	0.003447	0.007818	0.004099	0.005041	0.001824	0.002089	0.000914	...	0.002296	0.000985	0.000245	0.000000	0.000000	0.000000	0.000018	0.000587	0.000617	0.000000	...
а	0.000031	0.000744	0.000000	0.000000	0.000000	0.000000	0.002805	0.000000	0.000000	...	0.000018	0.000004	0.000000	0.000000	0.000000	0.000192	0.000143	0.000000	0.000000	0.000556	...
б	0.000681	0.007225	0.000106	0.000380	0.000149	0.000315	0.005095	0.000258	0.000053	...	0.000072	0.000207	0.000008	0.000000	0.000000	0.000006	0.000902	0.000000	0.000184	0.000975	...
в	0.003338	0.001715	0.000112	0.001368	0.000307	0.000198	0.004414	0.000043	0.000022	...	0.000352	0.000072	0.000000	0.000010	0.000219	0.002968	0.000002	0.000000	0.000225	0.003416	...
г	0.005379	0.005599	0.000035	0.002036	0.000018	0.000088	0.005763	0.000006	0.000016	...	0.000341	0.000012	0.000022	0.000012	0.001357	0.000509	0.000016	0.000004	0.000376	0.000000	...
д	0.006325	0.000033	0.000791	0.000666	0.001482	0.002022	0.000184	0.001676	0.000251	...	0.000012	0.000562	0.000321	0.000000	0.000000	0.000000	0.000008	0.000000	0.000734	0.000053	...
е	0.000016	0.000198	0.000000	0.000000	0.000000	0.000000	0.000057	0.000000	0.000000	...	0.000000	0.000000	0.000000	0.000000	0.000000	0.000020	0.000012	0.000000	0.000000	0.000000	...
...	0.002449	0.000374	0.000012	0.000206	0.000004	0.000016	0.000319	0.000008	0.000016	...	0.000002	0.000012	0.000002	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	...
ч	0.000325	0.000491	0.000000	0.000000	0.000000	0.000000	0.000004	0.000738	0.000000	...	0.000000	0.000002	0.000000	0.000000	0.000133	0.000000	0.000000	0.000000	0.000000	0.000002	...
ш	0.000431	0.002460	0.000002	0.000002	0.000000	0.000004	0.003702	0.000000	0.000000	...	0.000000	0.000151	0.000000	0.000000	0.000000	0.000229	0.000000	0.000000	0.000000	0.000000	...
щ	0.000059	0.000006	0.000000	0.000004	0.000000	0.000000	0.000000	0.002151	0.000000	...	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000038	0.000000	0.000000	0.000000	...
ь	0.000006	0.000025	0.000000	0.000000	0.000000	0.000000	0.000000	0.001366	0.000000	...	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000033	0.000000	0.000000	0.000000	...
ы	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000037	0.000000	0.000000	...	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000415	...
б	0.000408	0.000006	0.000159	0.000016	0.000104	0.000200	0.000018	0.000013	0.000063	...	0.000170	0.000417	0.000004	0.000000	0.000000	0.000000	0.000002	0.000000	0.000000	0.000006	...
э	0.011117	0.000008	0.000090	0.000108	0.000129	0.000055	0.000450	0.000004	0.000168	...	0.000088	0.000278	0.000029	0.000000	0.000000	0.000000	0.000008	0.000364	0.000485	0.000000	...
ю	0.000010	0.000000	0.000000	0.000000	0.000004	0.000000	0.000000	0.000000	0.000000	...	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000008	0.000000	0.000000	0.000000	...
я	0.002602	0.000000	0.000335	0.000022	0.000002	0.000327	0.000002	0.000016	0.000016	...	0.000110	0.000063	0.000162	0.000000	0.000000	0.000000	0.000000	0.000000	0.000045	0.000000	...
а	0.010490	0.000025	0.000051	0.000333	0.000084	0.000527	0.000100	0.000117	0.000227	...	0.000168	0.000016	0.000125	0.000000	0.000000	0.000000	0.000006	0.000000	0.000125	0.000080	...

[33 rows x 33 columns]

Sorted letter frequency without spaces:

(letter, count, frequency)

о	96681	0.11474726902425476
е	73380	0.08709213393531112
а	67115	0.0796564263977706
н	54835	0.06508172750535277
и	54644	0.06485503634179805
т	54553	0.06474703165130864
с	44600	0.05293416698712015
в	38984	0.04626873465977335
л	38734	0.045972018477110126
р	35246	0.041832234296592746
к	27833	0.03303400604826267
д	26983	0.032025171027207686
м	26495	0.03144598103864906
у	24993	0.02966331021320838
п	23121	0.027441499437426117
ь	19362	0.0229800749149018
я	18000	0.021363565151752525
ч	15249	0.018098500277726345
б	14655	0.017393502627718514
г	14232	0.01689145884665233
ы	13912	0.016511662132843396
з	12975	0.015399569880221611
ж	9615	0.011411704385227806
й	8437	0.01001357773251867
х	7171	0.008511006983512075
ш	6937	0.008233280636539293
ю	4733	0.005617430770180261
э	2971	0.0035261751147698194
щ	2521	0.0029920859859760064
ц	2336	0.0027725160108052166
ф	1049	0.0012450211024549111
ъ	204	0.00024212040505319527

H1 (entropy of single letters without spaces): 4.450720785464344
H2 (entropy of bigrams without overlapping without spaces): 4.127474285221844
H2 (entropy of bigrams with overlapping without spaces): 4.127369564159534

Bigram frequency matrix without overlapping (no spaces):																																										
	а	б	в	г	д	е	ж	з	и	...	ч	ш	щ	ц	б	ы	ь	э	ю	я	а	б	в	г	д	е	ж	з	и	...	ч	ш	щ	ц	б	ы	ь	э	ю	я		
а	0.000320	0.001306	0.005215	0.001016	0.003136	0.001961	0.002191	0.004524	0.001619	...	0.001823	0.001125	0.000313	0.000000	0.000000	0.000000	0.000000	0.000351	0.001037	0.002799	а	б	в	г	д	е	ж	з	и	...	ч	ш	щ	ц	б	ы	ь	э	ю	я		
б	0.000769	0.000005	0.000076	0.000000	0.000026	0.002535	0.000000	0.000021	0.000064	...	0.000045	0.000017	0.000204	0.000157	0.004527	0.000062	0.000097	0.000014	0.000021	б	ы	ь	э	ю	я	а	б	в	г	д	е	ж	з	и	...	ч	ш	щ	ц	б	ы	ь
в	0.006627	0.000216	0.000468	0.000247	0.001291	0.006115	0.000045	0.000793	0.003962	...	0.000406	0.000045	0.000000	0.000007	0.003169	0.000199	0.000316	0.000009	0.000068	в	г	д	е	ж	з	и	...	ч	ш	щ	ц	б	ы	ь	э	ю	я					
г	0.001222	0.000031	0.000150	0.000014	0.001401	0.000261	0.000007	0.000078	0.000698	...	0.000055	0.000000	0.000000	0.000000	0.000000	0.000000	0.000007	0.000002	0.000017	г	д	е	ж	з	и	...	ч	ш	щ	ц	б	ы	ь	э	ю	я						
д	0.006276	0.000059	0.001232	0.000026	0.000116	0.005120	0.000031	0.000059	0.002810	...	0.000088	0.000074	0.000002	0.000000	0.000532	0.001289	0.000019	0.000017	0.000513	д	е	ж	з	и	...	ч	ш	щ	ц	б	ы	ь	э	ю	я							
е	0.000287	0.002578	0.003715	0.004415	0.004299	0.002580	0.001170	0.002250	0.001588	...	0.001992	0.001268	0.001104	0.000000	0.000000	0.000000	0.000479	0.000275	0.000541	е	ж	з	и	...	ч	ш	щ	ц	б	ы	ь	э	ю	я								
ж	0.001552	0.000055	0.000052	0.000014	0.000069	0.005407	0.000009	0.000024	0.001536	...	0.000028	0.000000	0.000000	0.000000	0.000000	0.000057	0.000043	0.000000	0.000038	ж	з	и	...	ч	ш	щ	ц	б	ы	ь	э	ю	я									
з	0.005388	0.000176	0.001199	0.000411	0.000900	0.000249	0.000062	0.000059	0.000444	...	0.000052	0.000014	0.000000	0.000052	0.000320	0.000104	0.000026	0.000005	0.000529	з	и	...	ч	ш	щ	ц	б	ы	ь	э	ю	я										
и	0.000358	0.001707	0.005348	0.001021	0.003178	0.002431	0.000401	0.002554	0.001723	...	0.002471	0.000076	0.000154	0.000000	0.000000	0.000000	0.000233	0.000389	0.001683	и	...	ч	ш	щ	ц	б	ы	ь	э	ю	я											
й	0.000081	0.000221	0.000555	0.000242	0.000722	0.000114	0.000090	0.000157	0.000563	...	0.000577	0.000244	0.000007	0.000000	0.000000	0.000000	0.000102	0.000012	0.000090	й	...	ч	ш	щ	ц	б	ы	ь	э	ю	я											
к	0.000249	0.000010	0.000024	0.000005	0.000247	0.000563	0.000140	0.000116	0.000251	...	0.000263	0.000028	0.000005	0.000000	0.000000	0.000000	0.000007	0.000000	0.000126	к	...	ч	ш	щ	ц	б	ы	ь	э	ю	я											
л	0.000851	0.000304	0.000781	0.000294	0.000316	0.000481	0.000382	0.000199	0.000723	...	0.000036	0.000026	0.000005	0.000000	0.000098	0.000548	0.000107	0.000014	0.000167	л	...	ч	ш	щ	ц	б	ы	ь	э	ю	я											
м	0.003542	0.000287	0.000019	0.000266	0.000472	0.004968	0.000176	0.000259	0.000336	...	0.000387	0.000055	0.000005	0.000000	0.000047	0.000095	0.000119	0.000007	0.000565	м	...	ч	ш	щ	ц	б	ы	ь	э	ю	я											
н	0.012189	0.000247	0.000015	0.000000	0.000496	0.009496	0.012021	0.000024	0.000157	...	0.000373	0.000033	0.000126	0.000000	0.000241	0.001304	0.000064	0.000214	0.000224	н	...	ч	ш	щ	ц	б	ы	ь	э	ю	я											
о	0.000152	0.000564	0.001260	0.000538	0.001100	0.002958	0.000374	0.001056	0.002138	...	0.003022	0.001237	0.000294	0.000002	0.000000	0.000000	0.000477	0.000091	0.000130	о	...	ч	ш	щ	ц	б	ы	ь	э	ю	я											
п	0.000092	0.000005	0.000005	0.000000	0.000002	0.003107	0.000000	0.000000	0.000095	...	0.000021	0.000014	0.000000	0.000000	0.000261	0.000169	0.000000	0.000000	0.000750	п	...	ч	ш	щ	ц	б	ы	ь	э	ю	я											
р	0.000830	0.000169	0.000541	0.000211	0.000408	0.006112	0.000282	0.000069	0.006409	...	0.000003	0.000249	0.000021	0.000000	0.000095	0.001078	0.000005	0.000242	0.001132	р	...	ч	ш	щ	ц	б	ы	ь	э	ю	я											
с	0.002011	0.000247	0.002001	0.000114	0.000498	0.005369	0.000249	0.000107	0.001859	...	0.000582	0.000104	0.000002	0.000024	0.000263	0.003826	0.000003	0.000275	0.004194	с	...	ч	ш	щ	ц	б	ы	ь	э	ю	я											
т	0.007045	0.000451	0.003197	0.000990	0.000472	0.007238	0.000104	0.000190	0.004493	...	0.000700	0.000024	0.000028	0.000012	0.001652	0.007883	0.000126	0.000097	0.000574	т	...	ч	ш	щ	ц	б	ы	ь	э	ю	я											
у	0.000223	0.001097	0.001536	0.001840	0.002673	0.000344	0.002234	0.000508	0.000733	...	0.001424	0.000050	0.000313	0.000000	0.000000	0.000000	0.000121	0.000038	0.000237	у	...	ч	ш	щ	ц	б	ы	ь	э	ю	я											
ф	0.000266	0.000000	0.000000	0.000000	0.000002	0.000062	0.000000	0.000031	...	0.000002	0.000000	0.000000	0.000000	0.000000	0.000033	0.000109	0.000000	0.000000	0.000000	ф	...	ч	ш	щ	ц	б	ы	ь	э	ю	я											
х	0.000484	0.000116	0.000527	0.000059	0.000226	0.000475	0.000040	0.000066	0.000066	...	0.000005	0.000045	0.000000	0.000000	0.000000	0.000000	0.000062	0.000000	0.000038	х	...	ч	ш	щ	ц	б	ы	ь	э	ю	я											
ц	0.000653	0.000002	0.000078	0.000005	0.000009	0.000002	0.000000	0.000150	...	0.000017	0.000000	0.000000	0.000000	0.000185	0.000000	0.000012	0.000002	0.000009	0.000000	ц	...	ш	щ	ц	б	ы	ь	э	ю	я												
ч	0.002355	0.000019	0.000078	0.000009	0.000026	0.004555	0.000000	0.000007	0.001901	...	0.000007	0.000197	0.000000	0.000000	0.000000	0.000275	0.000007	0.000000	0.000012	ч	...	ш	щ	ц	б	ы	ь	э	ю	я												
ш	0.001085	0.000001	0.000007	0.000000	0.000000	0.002540	0.000000	0.000002	0.001609	...	0.000002	0.000000	0.000000	0.000000	0.000000	0.000073	0.000000	0.000000	0.000000	ш	...	щ	ц	б	ы	ь	э	ю	я													
щ	0.000451	0.000000	0.000000	0.000000	0.000000	0.001595	0.000000	0.000000	0.000629	...	0.000000	0.000000	0.000000	0.000000	0.000000	0.000055	0.000000	0.000000	0.000000	щ	...	ц	б	ы	ь	э	ю	я														
ц	0.000000	0.000000	0.000000	0.000000	0.000000	0.000038	0.000000	0.000000	0.000000	...	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	ц	...	б	ы	ь	э	ю	я															
б	0.000009	0.000453	0.001469	0.000178	0.000408	0.000971	0.000005	0.000237	0.000501	...	0.000358	0.000551	0.000012	0.000000	0.000000	0.000000	0.000131	0.000000	0.000057	б	...	ы	ь	э	ю	я																
ы	0.000304	0.000511	0.001460	0.000316	0.000560	0.000978	0.000095	0.000553	0.001422	...	0.000769	0.000399	0.000033	0.000000	0.000000	0.000000	0.000221	0.000061	0.000081	ы	...	ь	э	ю	я																	
ь	0.000000	0.000000	0.000007	0.000009	0.000007	0.000000	0.000000	0.000000	0.000000	...	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000021	0.000000	0.000000	ь	...	э	ю	я																		
э	0.000088	0.000446	0.000347	0.000062	0.000532	0.000005	0.000005	0.000058	0.001137	...	0.000252	0.000093	0.000209	0.000000	0.000000	0.000000	0.000057	0.000045	0.000065	э	...	ю	я																			
ю	0.000342	0.000461	0.001030	0.000292	0.001258	0.000463	0.000299	0.000558	0.000313	...	0.000068	0.000045	0.000142	0.000000	0.000000	0.000000	0.000202	0.000116	0.000244	ю	...	я																				

[32 rows x 32 columns]

The screenshot shows the 'Средняя информация' (Average Information) window. It contains several sections:

- Произвольная часть текста:** a_my_так_
- Использованные буквы:** (empty)
- Порядок n-граммы:** A list of n-grams from 5 to 50, each with a corresponding bar chart showing the frequency of the symbol 'а' in that context. The bar for '5' is highlighted in blue.
- Введенный символ:** (empty)
- Символ по счету:** (empty)
- Номер эксперимента:** 53
- Поле ввода символов:** (empty)
- Неравенство для энтропии:** $2.81985907951231 < H < 3.39117694773614$
- Двоичная таблица угаданных символов:** A table with 32 rows and 2 columns. The first column contains binary strings of length 32, and the second column contains the corresponding guessed symbols (mostly 'а' and 'б').
- Вероятности:** A list of 32 probabilities, $q[1]$ through $q[32]$, representing the estimated probabilities for each symbol in the alphabet.
- Строка состояния:** (empty)



