



Міністерство освіти і науки України
НТУУ «Київський політехнічний інститут»
Фізико-технічний інститут

КРИПТОГРАФІЯ

КОМП'ЮТЕРНИЙ ПРАКТИКУМ №1

Експериментальна оцінка ентропії на символ джерела
відкритого тексту.

Перевірив:
Чорний О. М.

Виконали:
студенти III курсу ФТІ
групи ФБ-71
Мельник Дмитрій
Безлюдный Вадим

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 1 Мб), 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.

Опис роботи та основні труднощі:

Для роботи було створено текст, який містив у собі цикл статей про шифрування та криптографію і роман М. А. Булгакова «Майстер і Маргарита», таким чином поєднавши науково-технічну та художню літературу. Текстовий файл розміром у 1,4 мб оброблюється програмою за 2-3 секунди в залежності від умов (з пробілами чи без, з перетином біграм чи без). Програма написана на мові Swift. Особливих труднощів під час роботи над комп'ютерним практикумом не виникло, за виключенням того, що декілька разів доводилось змінювати підхід до основного алгоритму програми задля збільшення її оптимізації. В найпершій версії текст проходився програмою $32+32*32$ рази, окремо для пошуку кожної монограми та кожної біграми, що займало набагато більше часу. Проблема була вирішена зміною алгоритму таким чином, що текст потрібно було пройти лише 1 раз, і створивши спеціальний словник та матрицю. З кожним кроком по тексут формується нова біграма (або через крок, якщо задано пошук біграм без перетину), а також перевіряється, чи наявна буква в словнику (якщо так, то її значення в словнику збільшується на 1). Сформовані біграми розбиваються на першу і другу, що відповідним чином вказує на i -тий рядок та j -тий стовпчик в матриці, що вказують на елемент, який також інкрементується. Далі кожне значення в словнику ділиться на суму всіх значень, а кожний елемент матриці – на суму всіх елементів. Таким чином ми отримуємо частоту кожної букви та біграми. Ентропії H_1 и H_2 рахуються за відповідною формулою.

Результати:

Монограми:

З пробілом

о – 0.0952772
е – 0.0684669
а – 0.068008
и – 0.0566651
н – 0.0531222
т – 0.0501816
с – 0.0469641
л – 0.040755
в – 0.0406411
р – 0.039265
к – 0.0288377
д – 0.0255413
м – 0.0249669
у – 0.0220353
п – 0.0216435
я – 0.0178317
г – 0.017157
ь – 0.0158502
ы – 0.0154723
з – 0.0148638
б – 0.0140168
ч – 0.0117437
й – 0.00886741
ж – 0.0085591
ш – 0.00743702
х – 0.00701942
ю – 0.00539621
ц – 0.00305019
э – 0.00245779
щ – 0.00240625
ф – 0.00180453

$H_1 = 4.3577$

Без пробіла

о – 0.113926
е – 0.0818684
а – 0.0813197
и – 0.0677566
н – 0.0635202
т – 0.060004
с – 0.0561567
л – 0.0487323
в – 0.048596
р – 0.0469506
к – 0.0344823
д – 0.0305407
м – 0.0298539
у – 0.0263485
п – 0.0258799
я – 0.021322
г – 0.0205153
ь – 0.0189527
ы – 0.0185008
з – 0.0177732
б – 0.0167604
ч – 0.0140424
й – 0.0106031
ж – 0.0102344
ш – 0.00889271
х – 0.00839337
ю – 0.00645244
ц – 0.00364723
э – 0.00293887
щ – 0.00287724
ф – 0.00215774

$H_1 = 4.45558$

Дані з Інтернету

ранг ♦	буква ♦	употреблений ♦	частотность ♦
1	о	55414481	10.97%
2	е	42691213	8.45%
3	а	40487008	8.01%
4	и	37153142	7.35%
5	н	33838881	6.70%
6	т	31620970	6.26%
7	с	27627040	5.47%
8	р	23916825	4.73%
9	в	22930719	4.54%
10	л	22230174	4.40%
11	к	17653469	3.49%
12	м	16203060	3.21%
13	д	15052118	2.98%
14	п	14201572	2.81%
15	у	13245712	2.62%
16	я	10139085	2.01%
17	ы	9595941	1.90%
18	ь	8784613	1.74%
19	г	8564640	1.70%
20	э	8329904	1.65%
21	б	8051767	1.59%
22	ч	7300193	1.44%
23	й	6106262	1.21%
24	х	4904176	0.97%
25	ж	4746916	0.94%
26	ш	3678738	0.73%
27	ю	3220715	0.64%
28	ц	2438807	0.48%
29	щ	1822476	0.36%
30	э	1610107	0.32%
31	ф	1335747	0.26%
32	ъ	185452	0.04%
33	ё	184928	0.04%

ё

Біграми:

Перетинаються, з пробілом – $H_1 = 3.92424$

Не перетинаються, з пробілом – $H_2 = 3.847661$

Перетинаються, без пробіла – $H_1 = 4.110573$

Не перетинаються, без пробіла – $H_2 = 4.107297$

Перетинаються, з пробілом

$N_1 = 3.92424$

	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	ы	ь	э	ю	я
а	0.000011	0.000884	0.005000	0.001189	0.002651	0.001434	0.001717	0.005659	0.000175	0.000940	0.005863	0.012550	0.003941	0.007246	0.000008	0.001338	0.004340	0.005138	0.006974	0.000196	0.000740	0.001397	0.000108	0.001094	0.001843	0.000352	0.000000	0.000000	0.000001	0.001076	0.000000
б	0.001422	0.000014	0.000122	0.000014	0.000030	0.003067	0.000009	0.000009	0.001055	0.000000	0.000307	0.001124	0.000058	0.000397	0.003231	0.000000	0.001642	0.000108	0.000009	0.001490	0.000000	0.000137	0.000004	0.000027	0.000005	0.000293	0.005524	0.000207	0.000000	0.000005	0.000001
в	0.008251	0.000018	0.000047	0.000050	0.000418	0.006363	0.000000	0.000712	0.004595	0.000000	0.000209	0.001550	0.001542	0.001668	0.010985	0.000341	0.000923	0.004395	0.000310	0.001182	0.000000	0.000056	0.000017	0.000056	0.001690	0.000006	0.003683	0.000247	0.000000	0.000003	0.000002
г	0.001347	0.000000	0.000057	0.000060	0.001442	0.000858	0.000001	0.000000	0.001059	0.000000	0.000110	0.002464	0.000005	0.000355	0.012885	0.000000	0.001724	0.000122	0.000010	0.000927	0.000000	0.000000	0.000002	0.000034	0.000010	0.000000	0.000000	0.000001	0.000000	0.000004	0.000000
д	0.006212	0.000045	0.001355	0.000009	0.000040	0.006445	0.000020	0.000004	0.003587	0.000000	0.000271	0.001141	0.000182	0.002380	0.005332	0.000141	0.002767	0.000568	0.000299	0.002358	0.000008	0.000060	0.000170	0.000046	0.000178	0.000000	0.000874	0.000805	0.000002	0.000061	0.000001
е	0.000046	0.001581	0.002216	0.004874	0.003797	0.002327	0.001481	0.001805	0.000262	0.003450	0.001591	0.008646	0.005712	0.011183	0.000770	0.001280	0.009709	0.006692	0.005513	0.000104	0.000021	0.001332	0.000339	0.001405	0.000973	0.000857	0.000000	0.000000	0.000000	0.000552	0.000000
ж	0.001950	0.000087	0.000000	0.000017	0.001073	0.005340	0.000026	0.000000	0.001948	0.000000	0.000116	0.000006	0.000009	0.001682	0.000074	0.000000	0.000004	0.000034	0.000000	0.000221	0.000000	0.000000	0.000000	0.000053	0.000000	0.000000	0.000000	0.000060	0.000000	0.000054	0.000000
з	0.007383	0.000245	0.001144	0.000597	0.002394	0.000361	0.000227	0.000105	0.000516	0.000000	0.000178	0.000303	0.000380	0.002347	0.000953	0.000028	0.000299	0.000325	0.000021	0.000396	0.000000	0.000000	0.000005	0.000020	0.000014	0.000000	0.000818	0.000930	0.000000	0.000097	0.000000
и	0.000129	0.000634	0.003564	0.000737	0.002158	0.004130	0.000562	0.004220	0.001336	0.001997	0.002814	0.006635	0.004267	0.004575	0.000280	0.000280	0.000798	0.004249	0.004673	0.000008	0.000050	0.002585	0.001703	0.001850	0.000493	0.000199	0.000000	0.000000	0.000000	0.000532	0.000000
й	0.000000	0.000008	0.000000	0.000001	0.000220	0.000005	0.000000	0.000000	0.000000	0.000000	0.000082	0.000114	0.000058	0.000597	0.000017	0.000001	0.000024	0.001029	0.000342	0.000000	0.000002	0.000002	0.000041	0.000081	0.000153	0.000000	0.000000	0.000000	0.000000	0.000000	0.000001
к	0.011978	0.000000	0.000521	0.000003	0.000010	0.000796	0.000030	0.000014	0.003787	0.000000	0.000021	0.000556	0.000008	0.002215	0.012095	0.000000	0.002183	0.000271	0.000685	0.002123	0.000000	0.000001	0.000016	0.000026	0.000012	0.000000	0.000000	0.000001	0.000000	0.000008	0.000000
л	0.009734	0.000031	0.000008	0.000177	0.000358	0.006255	0.000520	0.000006	0.009009	0.000000	0.000816	0.000147	0.000009	0.000402	0.008999	0.000243	0.000000	0.001720	0.000061	0.001707	0.000010	0.000008	0.000000	0.000274	0.000009	0.000003	0.001452	0.005295	0.000000	0.001540	0.000000
м	0.004035	0.000040	0.000001	0.000069	0.000001	0.005435	0.000000	0.000008	0.004757	0.000000	0.000093	0.000158	0.000032	0.001541	0.004995	0.000300	0.000065	0.000169	0.000063	0.003362	0.000009	0.000000	0.000034	0.000032	0.000008	0.000007	0.001164	0.000047	0.000000	0.000044	0.000000
н	0.015307	0.000019	0.000033	0.000175	0.001358	0.012881	0.000023	0.000019	0.011693	0.000000	0.000302	0.000006	0.000003	0.003959	0.012612	0.000008	0.000093	0.000835	0.000656	0.003187	0.000013	0.000009	0.001027	0.000311	0.000033	0.000163	0.004714	0.001447	0.000001	0.000200	0.000000
о	0.000009	0.004795	0.011769	0.006862	0.006792	0.003448	0.002478	0.001379	0.001185	0.004904	0.003936	0.009871	0.007232	0.009504	0.000291	0.001852	0.009164	0.010374	0.009888	0.000090	0.000486	0.000799	0.000104	0.001901	0.001571	0.000199	0.000000	0.000000	0.000034	0.001195	0.000000
п	0.001737	0.000000	0.000000	0.000005	0.000000	0.003689	0.000000	0.000000	0.001134	0.000000	0.000097	0.001051	0.000000	0.000097	0.012705	0.000071	0.009081	0.000008	0.000061	0.000953	0.000027	0.000000	0.000019	0.000122	0.000007	0.000002	0.000366	0.001255	0.000000	0.000000	0.000000
р	0.012421	0.000235	0.000434	0.000466	0.000489	0.008795	0.000466	0.000049	0.007462	0.000000	0.000341	0.000131	0.000448	0.000920	0.011370	0.000094	0.000028	0.000359	0.000856	0.004122	0.000024	0.000141	0.000119	0.0000100	0.0000409	0.000070	0.002235	0.001350	0.000001	0.000196	0.000000
с	0.002124	0.000076	0.002625	0.000029	0.000433	0.004987	0.000030	0.000040	0.002105	0.000000	0.007343	0.004044	0.001380	0.001296	0.004224	0.002458	0.000456	0.001333	0.021122	0.001138	0.000019	0.000240	0.000038	0.000476	0.000123	0.000000	0.000523	0.004122	0.000000	0.000164	0.000000
т	0.007964	0.000039	0.003755	0.000014	0.000229	0.006400	0.000001	0.000008	0.005067	0.000000	0.000594	0.000354	0.000036	0.001382	0.021716	0.000087	0.007212	0.001327	0.000171	0.002119	0.000011	0.000017	0.000157	0.000243	0.000001	0.000019	0.002005	0.007731	0.000002	0.000069	0.000000
у	0.000029	0.000833	0.001194	0.001726	0.002434	0.000297	0.001829	0.001297	0.000026	0.000097	0.001127	0.001744	0.001370	0.000399	0.000012	0.000957	0.000909	0.001957	0.001969	0.000005	0.000016	0.000470	0.000011	0.001119	0.001023	0.000383	0.000000	0.000000	0.000028	0.001405	0.000000
ф	0.000210	0.000000	0.000000	0.000000	0.000000	0.000204	0.000000	0.000000	0.000824	0.000000	0.000006	0.000098	0.000003	0.000000	0.000109	0.000000	0.000842	0.000018	0.000018	0.000115	0.000004	0.000000	0.000000	0.000000	0.000000	0.000000	0.000015	0.000004	0.000000	0.000003	0.000000
х	0.001193	0.000000	0.000174	0.000002	0.000001	0.000046	0.000000	0.000000	0.000222	0.000000	0.000000	0.000127	0.000074	0.000151	0.003093	0.000000	0.000148	0.000089	0.000051	0.000165	0.000000	0.000001	0.000000	0.000000	0.000011	0.000000	0.000000	0.000003	0.000000	0.000000	0.000000
ц	0.000802	0.000000	0.000030	0.000009	0.000001	0.001308	0.000000	0.000000	0.000353	0.000000	0.000181	0.000004	0.000002	0.000008	0.000530	0.000000	0.000000	0.000000	0.000002	0.000840	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000213	0.000000	0.000000	0.000000	0.000000
ч	0.003098	0.000000	0.000007	0.000000	0.000000	0.004540	0.000000	0.000000	0.002064	0.000000	0.000334	0.000053	0.000006	0.000756	0.000101	0.000000	0.000030	0.000000	0.005155	0.000870	0.000000	0.000000	0.000000	0.000000	0.000131	0.000000	0.000000	0.000255	0.000000	0.000000	0.000000
ш	0.002086	0.000000	0.000025	0.000000	0.000000	0.003270	0.000000	0.000000	0.002685	0.000000	0.000566	0.000604	0.000021	0.000532	0.000392	0.000053	0.000003	0.000000	0.000151	0.000432	0.000000	0.000000	0.000002	0.000000	0.000002	0.000000	0.000000	0.000376	0.000001	0.000000	0.000000
щ	0.000547	0.000000	0.000000	0.000000	0.000000	0.001840	0.000000	0.000000	0.001066	0.000000	0.000000	0.000000	0.000000	0.000043	0.000000	0.000000	0.000003	0.000000	0.000000	0.000130	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000021	0.000000	0.000000	0.000000
ы	0.000000	0.000530	0.001560	0.000119	0.000148	0.001583	0.000028	0.000080	0.000023	0.002036	0.000258	0.003389	0.002149	0.000366	0.000000	0.000154	0.000421	0.001123	0.000973	0.000006	0.000000	0.001522	0.000009	0.000246	0.000695	0.000004	0.000000	0.000000	0.000000	0.000000	0.000002
ь	0.000008	0.000139	0.000011	0.000045	0.000064	0.002124	0.000000	0.000191	0.000201	0.000000	0.001549	0.000000	0.000428	0.001552	0.000047	0.000001	0.000000	0.001188	0.000108	0.000000	0.000004	0.000001	0.000124	0.000079	0.000664	0.000012	0.000000	0.000000	0.000000	0.000735	0.000000
э	0.000000	0.000000	0.000000	0.000007	0.000005	0.000000	0.000000	0.000004	0.000000	0.000016	0.000075	0.000132	0.000007	0.000041	0.000000	0.000041	0.000026	0.000059	0.003296	0.000000	0.000002	0.000009	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
ю	0.000001	0.000588	0.000000	0.000003	0.000620	0.000001																									

Не перетинаються, з пробілом

$$H_2 = 3.847661$$

	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	ы	ь	э	ю	я
а	0.000012	0.001345	0.005367	0.001857	0.004535	0.001871	0.002435	0.006769	0.000289	0.001518	0.011494	0.014884	0.006714	0.007539	0.000016	0.002775	0.006270	0.007574	0.008589	0.000232	0.000149	0.002051	0.000140	0.001496	0.003905	0.000626	0.000000	0.000000	0.000003	0.001336	0.000000
б	0.000980	0.000026	0.000104	0.000025	0.000067	0.001825	0.000020	0.000020	0.000751	0.000000	0.000535	0.000590	0.000124	0.000653	0.001628	0.000000	0.001574	0.000175	0.000021	0.000959	0.000000	0.000298	0.000001	0.000009	0.000005	0.000642	0.001575	0.000399	0.000000	0.000000	0.000000
в	0.008083	0.000003	0.000004	0.000066	0.000210	0.003956	0.000000	0.000038	0.004938	0.000000	0.000102	0.001850	0.000018	0.001845	0.010139	0.000045	0.000356	0.000780	0.000068	0.001395	0.000000	0.000001	0.000033	0.000003	0.001459	0.000008	0.001733	0.000255	0.000000	0.000000	0.000000
г	0.001151	0.000000	0.000016	0.000003	0.000722	0.000246	0.000000	0.000000	0.000991	0.000000	0.000070	0.000877	0.000001	0.000472	0.012072	0.000000	0.000833	0.000052	0.000003	0.000456	0.000000	0.000000	0.000005	0.000002	0.000012	0.000000	0.000002	0.000000	0.000000	0.000003	0.000000
д	0.005741	0.000008	0.000388	0.000000	0.000002	0.004135	0.000003	0.000002	0.003945	0.000000	0.000099	0.000423	0.000131	0.003218	0.002226	0.000231	0.001642	0.000823	0.000009	0.001707	0.000000	0.000003	0.000284	0.000023	0.000183	0.000000	0.000607	0.000762	0.000004	0.000004	0.000000
е	0.000092	0.003068	0.003001	0.004571	0.003875	0.002276	0.001715	0.002682	0.000394	0.004271	0.002695	0.014399	0.006586	0.014438	0.001586	0.002411	0.014247	0.008447	0.007988	0.000223	0.000024	0.001322	0.000398	0.002089	0.001279	0.000444	0.000000	0.000000	0.000000	0.000653	0.000000
ж	0.001516	0.000192	0.000000	0.000003	0.000971	0.003966	0.000006	0.000000	0.001603	0.000000	0.000131	0.000000	0.000001	0.001348	0.000074	0.000000	0.000000	0.000007	0.000000	0.000203	0.000000	0.000000	0.000000	0.000001	0.000000	0.000000	0.000000	0.000014	0.000000	0.000003	0.000000
з	0.005727	0.000264	0.000917	0.000644	0.003905	0.000160	0.000193	0.000149	0.000277	0.000000	0.000217	0.000171	0.000361	0.001070	0.000361	0.000060	0.000282	0.000014	0.000027	0.000226	0.000000	0.000000	0.000000	0.000004	0.000013	0.000000	0.000644	0.000068	0.000000	0.000003	0.000000
и	0.000151	0.000481	0.003508	0.000455	0.002878	0.004585	0.000371	0.001769	0.001898	0.002580	0.003915	0.008930	0.005242	0.004830	0.000170	0.000242	0.001056	0.004996	0.005845	0.000002	0.000092	0.002908	0.002245	0.002263	0.000580	0.000271	0.000000	0.000000	0.000000	0.000698	0.000000
й	0.000000	0.000006	0.000001	0.000000	0.000109	0.000000	0.000000	0.000001	0.000000	0.000000	0.000046	0.000014	0.000027	0.000557	0.000010	0.000000	0.000004	0.000584	0.000289	0.000000	0.000000	0.000003	0.000034	0.000005	0.000108	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
к	0.008354	0.000001	0.000989	0.000004	0.000023	0.000586	0.000000	0.000032	0.003452	0.000000	0.000036	0.000424	0.000008	0.000743	0.007211	0.000001	0.001039	0.000377	0.000278	0.001283	0.000001	0.000000	0.000027	0.000056	0.000018	0.000000	0.000000	0.000000	0.000000	0.000002	0.000000
л	0.010643	0.000027	0.000003	0.000024	0.000009	0.006412	0.000486	0.000006	0.008525	0.000000	0.000267	0.000068	0.000005	0.000100	0.005874	0.000202	0.000000	0.001419	0.000012	0.001981	0.000002	0.000000	0.000000	0.000089	0.000000	0.000000	0.001960	0.003727	0.000001	0.000872	0.000000
м	0.001523	0.000044	0.000002	0.000006	0.000002	0.006944	0.000000	0.000000	0.003363	0.000000	0.000125	0.000122	0.000006	0.000234	0.002414	0.000425	0.000035	0.000196	0.000037	0.003347	0.000004	0.000001	0.000006	0.000022	0.000006	0.000013	0.000630	0.000006	0.000001	0.000027	0.000000
н	0.011754	0.000020	0.000016	0.000338	0.001442	0.004717	0.000016	0.000029	0.011227	0.000000	0.000200	0.000001	0.000001	0.004639	0.011570	0.000017	0.000009	0.001109	0.000613	0.002598	0.000018	0.000002	0.001838	0.000125	0.000014	0.000013	0.005411	0.001469	0.000001	0.000262	0.000000
о	0.000020	0.002950	0.016254	0.010349	0.009254	0.003812	0.003700	0.002504	0.000517	0.006783	0.007024	0.018016	0.011168	0.005235	0.000555	0.0001804	0.015375	0.015789	0.010653	0.000157	0.000273	0.001241	0.000168	0.002602	0.002767	0.000171	0.000000	0.000000	0.000069	0.001403	0.000000
п	0.001560	0.000002	0.000000	0.000000	0.000000	0.001543	0.000000	0.000000	0.000844	0.000000	0.000130	0.000367	0.000000	0.000107	0.003076	0.000122	0.002285	0.000013	0.000074	0.000729	0.000004	0.000000	0.000001	0.000181	0.000007	0.000000	0.000235	0.000080	0.000000	0.000000	0.000000
р	0.011609	0.000027	0.000124	0.000595	0.000453	0.010091	0.000432	0.000018	0.010521	0.000000	0.000207	0.000161	0.000733	0.000363	0.011117	0.000012	0.000017	0.000203	0.000680	0.004815	0.000025	0.000146	0.000107	0.000093	0.000225	0.000053	0.001613	0.000285	0.000002	0.000109	0.000000
с	0.000714	0.000006	0.000373	0.000018	0.000026	0.006259	0.000003	0.000001	0.001662	0.000000	0.006404	0.001583	0.000363	0.001153	0.001207	0.001690	0.000070	0.000262	0.017119	0.000439	0.000006	0.000148	0.000009	0.000104	0.000012	0.000000	0.000342	0.003096	0.000001	0.000199	0.000000
т	0.007174	0.000086	0.004254	0.000024	0.000501	0.004516	0.000002	0.000019	0.005936	0.000000	0.000891	0.000425	0.000043	0.001575	0.026299	0.000192	0.012804	0.001641	0.000360	0.001989	0.000016	0.000036	0.000341	0.000307	0.000002	0.000001	0.001666	0.006748	0.000001	0.000035	0.000000
у	0.000041	0.000863	0.001577	0.001066	0.003504	0.000159	0.001719	0.002387	0.000050	0.000125	0.001521	0.001659	0.002040	0.000587	0.000026	0.000584	0.001229	0.002088	0.003075	0.000007	0.000030	0.000600	0.000017	0.001318	0.001322	0.000548	0.000000	0.000000	0.000062	0.001567	0.000000
ф	0.000260	0.000000	0.000000	0.000000	0.000000	0.000138	0.000000	0.000000	0.001428	0.000000	0.000001	0.000000	0.000002	0.000001	0.000066	0.000000	0.000065	0.000011	0.000002	0.000142	0.000008	0.000000	0.000000	0.000001	0.000000	0.000000	0.000009	0.000003	0.000000	0.000000	0.000000
х	0.001029	0.000000	0.000176	0.000002	0.000000	0.000033	0.000000	0.000000	0.000184	0.000000	0.000000	0.000023	0.000004	0.000069	0.001456	0.000001	0.000046	0.000109	0.000013	0.000052	0.000001	0.000000	0.000000	0.000001	0.000012	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
ц	0.000466	0.000000	0.000006	0.000000	0.000001	0.001299	0.000001	0.000000	0.000278	0.000000	0.000229	0.000008	0.000001	0.000002	0.000311	0.000000	0.000000	0.000000	0.000002	0.000208	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000248	0.000000	0.000000	0.000000	0.000000
ч	0.003579	0.000000	0.000000	0.000000	0.000000	0.002838	0.000000	0.000000	0.002169	0.000000	0.000357	0.000022	0.000010	0.000827	0.000093	0.000000	0.000025	0.000000	0.000498	0.000215	0.000000	0.000000	0.000000	0.000000	0.000019	0.000000	0.000000	0.000076	0.000000	0.000000	0.000000
ш	0.001715	0.000000	0.000001	0.000000	0.000000	0.002367	0.000000	0.000000	0.002567	0.000000	0.000435	0.000408	0.000003	0.000553	0.000026	0.000012	0.000002	0.000001	0.000019	0.000256	0.000000	0.000000	0.000002	0.000000	0.000002	0.000000	0.000000	0.000319	0.000000	0.000001	0.000000
щ	0.000493	0.000000	0.000000	0.000000	0.000000	0.001989	0.000000	0.000000	0.001075	0.000000	0.000000	0.000000	0.000000	0.000011	0.000001	0.000000	0.000008	0.000000	0.000000	0.000142	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000002	0.000000	0.000000	0.000000
ы	0.000000	0.000262	0.002520	0.000224	0.000229	0.002290	0.000056	0.000149	0.000052	0.002572	0.000229	0.007099	0.002841	0.000751	0.000000	0.000289	0.000913	0.002199	0.001862	0.000012	0.000001	0.002077	0.000021	0.000259	0.000708	0.000006	0.000000	0.000000	0.000000	0.000000	0.000000
ь	0.000000	0.000203	0.000024	0.000093	0.000102	0.003972	0.000001	0.000384	0.000312	0.000001	0.002627	0.000001	0.000881	0.002274	0.000089	0.000002	0.000000	0.001722	0.000163	0.000001	0.000005	0.000000	0.000160	0.000172	0.001391	0.000020	0.000000	0.000000	0.000000	0.000807	0.000000
э	0.000000	0.000000	0.000000	0.000001	0.000000	0.000000	0.000000	0.000001	0.000000	0.000007	0.000003	0.000005	0.000004	0.000002	0.000000	0.000000	0.000005	0.000001	0.000173	0.000000	0.000002	0.000001	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
ю	0.000000	0.001180	0.000000	0.000000	0.001216	0.000001																									

Перетинаються, без пробіла

N₁ = 4.110573

	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	ы	ь	э	ю	я
а	0.000283	0.001501	0.006942	0.001484	0.002992	0.001723	0.001524	0.004951	0.001557	0.000755	0.006017	0.010353	0.003931	0.007897	0.001447	0.003035	0.003994	0.006178	0.006445	0.000567	0.000707	0.001232	0.000137	0.001372	0.001562	0.000289	0.000000	0.000000	0.000334	0.000867	0.000000
б	0.001153	0.000016	0.000108	0.000013	0.000028	0.002479	0.000008	0.000011	0.000889	0.000000	0.000249	0.000901	0.000049	0.000330	0.002703	0.000007	0.001318	0.000097	0.000009	0.001231	0.000000	0.000110	0.000003	0.000024	0.000004	0.000234	0.004427	0.000166	0.000041	0.000004	0.000000
в	0.006737	0.000275	0.000596	0.000502	0.000757	0.005283	0.000060	0.000716	0.005034	0.000000	0.000851	0.001410	0.001557	0.001929	0.009316	0.001126	0.001437	0.004609	0.000823	0.001089	0.000034	0.000088	0.000042	0.000231	0.001411	0.000009	0.002952	0.000198	0.000304	0.000004	0.000000
г	0.001105	0.000081	0.000293	0.000086	0.001264	0.000716	0.000017	0.000064	0.001002	0.000000	0.000201	0.001992	0.000047	0.000481	0.010479	0.000261	0.001428	0.000331	0.000063	0.000789	0.000011	0.000009	0.000005	0.000168	0.000012	0.000000	0.000000	0.000001	0.000031	0.000003	0.000000
д	0.004996	0.000081	0.001205	0.001120	0.000076	0.005190	0.000024	0.000050	0.002984	0.000000	0.000318	0.000930	0.000184	0.002056	0.004341	0.000238	0.002259	0.000611	0.000280	0.001915	0.000022	0.000051	0.000140	0.000070	0.000149	0.000000	0.000700	0.000645	0.000026	0.000050	0.000000
е	0.000217	0.002289	0.003777	0.004472	0.003995	0.002269	0.001414	0.002056	0.001557	0.002766	0.002256	0.007325	0.005633	0.010537	0.001992	0.003084	0.008305	0.007624	0.005259	0.000610	0.000094	0.001254	0.000304	0.001663	0.000859	0.000690	0.000000	0.000000	0.000321	0.000445	0.000000
ж	0.001564	0.000079	0.000022	0.000019	0.000872	0.004289	0.000023	0.000008	0.001572	0.000000	0.000105	0.000008	0.000022	0.001377	0.000071	0.000013	0.000005	0.000042	0.000019	0.000180	0.000000	0.000001	0.000000	0.000047	0.000000	0.000000	0.000000	0.000048	0.000012	0.000044	0.000000
з	0.005934	0.000238	0.001032	0.000523	0.001998	0.000311	0.000193	0.000118	0.000459	0.000000	0.000281	0.000267	0.000387	0.001998	0.000850	0.000210	0.000290	0.000440	0.000089	0.000347	0.000007	0.000006	0.000008	0.000040	0.000023	0.000001	0.000655	0.000745	0.000042	0.000078	0.000000
и	0.000343	0.001454	0.005259	0.001265	0.002746	0.003833	0.000643	0.003945	0.002783	0.001603	0.003468	0.005668	0.004063	0.005766	0.001908	0.002615	0.001268	0.005677	0.004718	0.000533	0.000177	0.002220	0.001413	0.002085	0.000532	0.000169	0.000000	0.000001	0.000257	0.000432	0.000000
й	0.000188	0.000336	0.000789	0.000339	0.000695	0.000119	0.000135	0.000159	0.000629	0.000002	0.000675	0.000239	0.000401	0.001078	0.000538	0.000811	0.000359	0.001750	0.000536	0.000176	0.000078	0.000069	0.000065	0.000322	0.000194	0.000003	0.000000	0.000000	0.000067	0.000004	0.000000
к	0.009665	0.000463	0.000800	0.000155	0.000247	0.000718	0.000217	0.000099	0.003358	0.000000	0.000380	0.000501	0.000196	0.002338	0.010090	0.000397	0.001853	0.000709	0.000794	0.001790	0.000025	0.000041	0.000025	0.000158	0.000031	0.000002	0.000000	0.000001	0.000104	0.000008	0.000000
л	0.007915	0.000308	0.000834	0.000387	0.000639	0.005476	0.000463	0.000154	0.007695	0.000001	0.001305	0.000192	0.000140	0.001071	0.008161	0.000983	0.000245	0.002286	0.000354	0.001517	0.000048	0.000039	0.000011	0.000606	0.000053	0.000006	0.001164	0.004244	0.000118	0.001235	0.000000
м	0.003344	0.000350	0.000827	0.000387	0.000403	0.004493	0.000129	0.000172	0.004576	0.000000	0.000649	0.000300	0.000266	0.001877	0.004674	0.001064	0.000253	0.001101	0.000323	0.002894	0.000061	0.000072	0.000041	0.000504	0.000067	0.000008	0.000933	0.000037	0.000078	0.000037	0.000000
н	0.012319	0.000327	0.000562	0.000282	0.001274	0.010397	0.000054	0.000179	0.009726	0.000000	0.000451	0.000046	0.000118	0.003713	0.010416	0.000585	0.000186	0.001209	0.000710	0.002731	0.000031	0.000057	0.000828	0.000350	0.000042	0.000132	0.003778	0.001160	0.000048	0.000161	0.000000
о	0.000210	0.005222	0.012358	0.006115	0.006548	0.003404	0.002400	0.001656	0.002480	0.003931	0.004376	0.008301	0.006584	0.010075	0.002382	0.003803	0.007912	0.010760	0.009181	0.000631	0.000514	0.000809	0.000125	0.002750	0.001345	0.000166	0.000000	0.000000	0.000386	0.000963	0.000000
п	0.001393	0.000001	0.000003	0.000006	0.000001	0.002957	0.000000	0.000000	0.000915	0.000000	0.000082	0.000843	0.000001	0.000083	0.010185	0.000060	0.007278	0.000013	0.000050	0.000764	0.000023	0.000000	0.000015	0.000099	0.000006	0.000002	0.000294	0.001006	0.000002	0.000000	0.000000
р	0.010007	0.000251	0.000978	0.000408	0.000448	0.007065	0.000382	0.000080	0.006114	0.000000	0.000371	0.000119	0.000397	0.000900	0.009224	0.000280	0.000060	0.001805	0.000733	0.003352	0.000038	0.000125	0.000098	0.000111	0.000337	0.000056	0.001791	0.001082	0.000022	0.000157	0.000000
с	0.001744	0.000198	0.002320	0.000159	0.000527	0.004048	0.000125	0.000105	0.001798	0.000000	0.006125	0.003310	0.001196	0.001458	0.003575	0.002263	0.000477	0.001425	0.017209	0.001045	0.000048	0.000207	0.000041	0.000422	0.000118	0.000001	0.000419	0.003303	0.000059	0.000134	0.000000
т	0.006461	0.000321	0.003639	0.000158	0.000462	0.005339	0.000112	0.000125	0.004418	0.000000	0.000759	0.000360	0.000206	0.001623	0.017774	0.000520	0.005893	0.001613	0.000413	0.001802	0.000035	0.000051	0.000136	0.000407	0.000016	0.000018	0.001606	0.006196	0.000122	0.000056	0.000000
у	0.000148	0.000890	0.001580	0.001586	0.002261	0.000352	0.001545	0.001182	0.000664	0.000078	0.001425	0.001498	0.001370	0.001084	0.000423	0.001391	0.000894	0.002242	0.001847	0.000123	0.000047	0.000422	0.000020	0.001318	0.000842	0.000311	0.000000	0.000000	0.000102	0.001127	0.000000
ф	0.000174	0.000009	0.000022	0.000004	0.000003	0.000165	0.000001	0.000005	0.000688	0.000000	0.000012	0.000081	0.000007	0.000015	0.0000105	0.000020	0.000695	0.000036	0.000023	0.000096	0.000004	0.000001	0.000000	0.000002	0.000002	0.000000	0.000012	0.000003	0.000000	0.000002	0.000000
х	0.000999	0.000158	0.000530	0.000149	0.000212	0.000102	0.000048	0.000083	0.000525	0.000000	0.000234	0.000293	0.000201	0.000421	0.002750	0.000450	0.000242	0.000544	0.000161	0.000235	0.000027	0.000017	0.000012	0.000076	0.000028	0.000001	0.000000	0.000003	0.000039	0.000000	0.000000
ц	0.000648	0.000010	0.000053	0.000019	0.000009	0.001055	0.000001	0.000005	0.000307	0.000000	0.000163	0.000006	0.000011	0.000028	0.000445	0.000035	0.000007	0.000036	0.000015	0.000675	0.000001	0.000001	0.000000	0.000005	0.000001	0.000000	0.000171	0.000000	0.000004	0.000000	0.000000
ч	0.002488	0.000011	0.000041	0.000152	0.000008	0.003641	0.000001	0.000003	0.001668	0.000000	0.000278	0.000046	0.000010	0.000622	0.000100	0.000034	0.000042	0.000028	0.004138	0.000703	0.000000	0.000002	0.000000	0.000007	0.000107	0.000000	0.000000	0.000204	0.000002	0.000000	0.000000
ш	0.001673	0.000010	0.000028	0.000003	0.000006	0.002621	0.000001	0.000002	0.002157	0.000000	0.000460	0.000486	0.000022	0.000435	0.000320	0.000055	0.000007	0.000008	0.000124	0.000347	0.000000	0.000000	0.000002	0.000001	0.000002	0.000000	0.000000	0.000302	0.000003	0.000000	0.000000
щ	0.000438	0.000000	0.000001	0.000000	0.000000	0.001475	0.000000	0.000000	0.000856	0.000000	0.000001	0.000000	0.000000	0.000035	0.000000	0.000000	0.000004	0.000000	0.000000	0.000104	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000017	0.000000	0.000000	0.000000
ы	0.000054	0.000684	0.001713	0.000211	0.000341	0.001376	0.000064	0.000210	0.000487	0.001633	0.000444	0.002773	0.001942	0.000848	0.000308	0.000619	0.000454	0.001376	0.000981	0.000098	0.000026	0.001261	0.000015	0.000297	0.000576	0.000006	0.000000	0.000000	0.000056	0.000000	0.000000
ь	0.000473	0.000455	0.001302	0.000284	0.000457	0.002190	0.000071	0.000418	0.001009	0.000001	0.001960	0.000157	0.000632	0.002258	0.000776	0.000984	0.000235	0.002096	0.000553	0.000207	0.000041	0.000082	0.000113	0.000467	0.000568	0.000011	0.000000	0.000000	0.000201	0.000591	0.000000
э	0.000000	0.000000	0.000003	0.000006	0.000007	0.000000	0.000000	0.000003	0.000000	0.000013	0.000061	0.000106	0.000006	0.000033	0.000000	0.000034	0.000020	0.000048	0.002642	0.000000	0.000001	0.000007	0.000001	0.000001	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
ю	0.000121	0.000576	0.000330	0.000140	0.000677	0.000085																									

Не перетинаються, без пробіла

$$H_2 = 4.107297$$

	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	ы	ь	э	ю	я
а	0.000292	0.001569	0.006933	0.001587	0.003126	0.001777	0.001512	0.004951	0.001612	0.000777	0.006482	0.010556	0.003964	0.007649	0.001479	0.003226	0.003818	0.006520	0.006606	0.000608	0.000684	0.001305	0.000134	0.001367	0.001670	0.000305	0.000000	0.000000	0.000367	0.000899	0.000000
б	0.001120	0.000011	0.000104	0.000013	0.000029	0.002420	0.000011	0.000012	0.000908	0.000000	0.000266	0.000882	0.000059	0.000333	0.002698	0.000008	0.001311	0.000092	0.000009	0.001212	0.000000	0.000107	0.000002	0.000019	0.000003	0.000265	0.004403	0.000176	0.000047	0.000005	0.000000
в	0.006590	0.000253	0.000576	0.000477	0.000734	0.005253	0.000059	0.000672	0.004986	0.000000	0.000801	0.001368	0.001598	0.001942	0.009196	0.001045	0.001398	0.004321	0.000754	0.001115	0.000032	0.000091	0.000046	0.000209	0.001263	0.000006	0.002922	0.000214	0.000300	0.000006	0.000000
г	0.001093	0.000094	0.000296	0.000079	0.001271	0.000682	0.000015	0.000057	0.000906	0.000000	0.000180	0.001907	0.000044	0.000495	0.010220	0.000245	0.001372	0.000344	0.000059	0.000748	0.000009	0.000009	0.000006	0.000144	0.000015	0.000000	0.000001	0.000001	0.000034	0.000003	0.000000
д	0.004963	0.000072	0.001163	0.001147	0.000072	0.005046	0.000019	0.000049	0.002999	0.000000	0.000285	0.000853	0.000147	0.002037	0.004210	0.000247	0.001697	0.000612	0.000292	0.001848	0.000019	0.000046	0.000146	0.000059	0.000149	0.000001	0.000528	0.000626	0.000024	0.000009	0.000000
е	0.000237	0.002424	0.003874	0.004657	0.004010	0.002299	0.001424	0.002101	0.001539	0.002371	0.002316	0.007900	0.005664	0.010660	0.002055	0.003141	0.008597	0.007846	0.005281	0.000630	0.000094	0.001278	0.000319	0.001682	0.000886	0.000654	0.000000	0.000000	0.000307	0.000380	0.000000
ж	0.001555	0.000089	0.000019	0.000020	0.000832	0.004205	0.000020	0.000003	0.001564	0.000000	0.000109	0.000011	0.000022	0.001026	0.000076	0.000016	0.000004	0.000044	0.000022	0.000181	0.000000	0.000001	0.000000	0.000050	0.000000	0.000000	0.000000	0.000039	0.000013	0.000046	0.000000
з	0.005781	0.000270	0.001041	0.000553	0.001976	0.000271	0.000194	0.000127	0.000439	0.000000	0.000298	0.000263	0.000380	0.001899	0.000831	0.000218	0.000296	0.000462	0.000089	0.000339	0.000007	0.000006	0.000008	0.000043	0.000013	0.000001	0.000611	0.000113	0.000046	0.000003	0.000000
и	0.000335	0.001443	0.005215	0.001273	0.002746	0.003934	0.000643	0.003878	0.002918	0.001709	0.003502	0.005938	0.004069	0.005731	0.001859	0.002557	0.001232	0.005861	0.004721	0.000518	0.000181	0.002257	0.001449	0.002098	0.000510	0.000168	0.000000	0.000000	0.000257	0.000442	0.000000
й	0.000184	0.000345	0.000853	0.000333	0.000715	0.000131	0.000151	0.000161	0.000656	0.000003	0.000682	0.000262	0.000415	0.001104	0.000576	0.000869	0.000355	0.001773	0.000550	0.000190	0.000084	0.000069	0.000064	0.000342	0.000194	0.000003	0.000000	0.000000	0.000064	0.000006	0.000000
к	0.009738	0.000431	0.000761	0.000144	0.000196	0.000738	0.000207	0.000082	0.003359	0.000000	0.000329	0.000484	0.000171	0.002204	0.009885	0.000365	0.001759	0.000668	0.000758	0.001748	0.000023	0.000043	0.000020	0.000142	0.000021	0.000001	0.000000	0.000001	0.000093	0.000005	0.000000
л	0.007785	0.000303	0.000764	0.000378	0.000618	0.005558	0.000444	0.000138	0.007420	0.000002	0.001300	0.000187	0.000135	0.001003	0.007955	0.000932	0.000235	0.002033	0.000330	0.001506	0.000046	0.000044	0.000011	0.000583	0.000054	0.000006	0.001169	0.004089	0.000112	0.001205	0.000000
м	0.003403	0.000330	0.000818	0.000375	0.000415	0.004382	0.000114	0.000169	0.004359	0.000001	0.000643	0.000293	0.000270	0.001841	0.004552	0.001079	0.000253	0.001067	0.000324	0.002971	0.000062	0.000069	0.000042	0.000513	0.000062	0.000004	0.000912	0.000034	0.000069	0.000040	0.000000
н	0.011950	0.000322	0.000584	0.000283	0.001803	0.010148	0.000048	0.000187	0.009749	0.000000	0.000453	0.000052	0.000123	0.003580	0.010609	0.000618	0.000183	0.001199	0.000741	0.002575	0.000031	0.000065	0.000835	0.000339	0.000042	0.000114	0.003951	0.001184	0.000054	0.000169	0.000000
о	0.000213	0.005014	0.012553	0.006410	0.006722	0.003420	0.002449	0.001729	0.002488	0.003990	0.004418	0.008603	0.006817	0.010144	0.002448	0.003858	0.008236	0.011044	0.009177	0.000643	0.000535	0.000818	0.000125	0.002945	0.001403	0.000161	0.000000	0.000000	0.000409	0.000961	0.000000
п	0.001384	0.000001	0.000003	0.000006	0.000002	0.002881	0.000000	0.000000	0.000899	0.000000	0.000076	0.000801	0.000002	0.000085	0.009790	0.000073	0.006935	0.000016	0.000051	0.000744	0.000019	0.000000	0.000015	0.000093	0.000006	0.000001	0.000300	0.000982	0.000003	0.000000	0.000000
р	0.010115	0.000247	0.001016	0.000412	0.000480	0.007639	0.000372	0.000085	0.006316	0.000000	0.000329	0.000124	0.000420	0.000863	0.009464	0.000287	0.000059	0.001716	0.000722	0.003416	0.000043	0.000124	0.000087	0.000104	0.000339	0.000052	0.001800	0.001232	0.000022	0.000144	0.000000
с	0.001690	0.000197	0.002183	0.000148	0.000502	0.004182	0.000115	0.000099	0.001720	0.000000	0.006078	0.003186	0.001163	0.001276	0.003479	0.002225	0.000428	0.001344	0.017086	0.001006	0.000038	0.000204	0.000041	0.000401	0.000117	0.000000	0.000407	0.002585	0.000049	0.000151	0.000000
т	0.006463	0.000331	0.003696	0.000158	0.000467	0.004974	0.000112	0.000131	0.004419	0.000000	0.000790	0.000342	0.000216	0.001439	0.018123	0.000507	0.006003	0.001690	0.000418	0.001830	0.000034	0.000049	0.000144	0.000403	0.000017	0.000016	0.001561	0.005657	0.000125	0.000048	0.000000
у	0.000143	0.000861	0.001587	0.001523	0.002313	0.000347	0.001549	0.001221	0.000672	0.000069	0.001415	0.001594	0.001424	0.001111	0.000407	0.001380	0.000929	0.002272	0.001936	0.000125	0.000046	0.000423	0.000019	0.001349	0.000892	0.000319	0.000000	0.000000	0.000101	0.001152	0.000000
ф	0.000157	0.000012	0.000026	0.000006	0.000004	0.000170	0.000001	0.000006	0.000690	0.000000	0.000013	0.000084	0.000008	0.000014	0.000102	0.000019	0.000678	0.000037	0.000020	0.000105	0.000004	0.000000	0.000000	0.000001	0.000002	0.000000	0.000008	0.000003	0.000000	0.000002	0.000000
х	0.001014	0.000150	0.000494	0.000132	0.000204	0.000102	0.000046	0.000074	0.000491	0.000000	0.000215	0.000280	0.000183	0.000399	0.002638	0.000440	0.000240	0.000522	0.000154	0.000240	0.000028	0.000019	0.000009	0.000072	0.000029	0.000001	0.000000	0.000000	0.000042	0.000000	0.000000
ц	0.000629	0.000010	0.000052	0.000019	0.000007	0.001029	0.000000	0.000003	0.000293	0.000000	0.000168	0.000006	0.000009	0.000029	0.000452	0.000029	0.000006	0.000037	0.000012	0.000674	0.000001	0.000002	0.000000	0.000005	0.000000	0.000000	0.000177	0.000000	0.000006	0.000000	0.000000
ч	0.002543	0.000013	0.000035	0.000159	0.000007	0.003614	0.000001	0.000003	0.001660	0.000000	0.000266	0.000040	0.000013	0.000620	0.000087	0.000025	0.000040	0.000026	0.003802	0.000705	0.000000	0.000002	0.000000	0.000003	0.000108	0.000000	0.000000	0.000205	0.000003	0.000000	0.000000
ш	0.001633	0.000009	0.000026	0.000001	0.000006	0.002649	0.000000	0.000002	0.002244	0.000000	0.000401	0.000487	0.000024	0.000442	0.000319	0.000061	0.000007	0.000006	0.000123	0.000350	0.000000	0.000000	0.000003	0.000002	0.000001	0.000000	0.000000	0.000303	0.000002	0.000001	0.000000
щ	0.000415	0.000000	0.000001	0.000000	0.000000	0.001427	0.000000	0.000000	0.000818	0.000000	0.000002	0.000000	0.000000	0.000026	0.000000	0.000000	0.000003	0.000001	0.000000	0.000103	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000011	0.000000	0.000000	0.000000
ы	0.000052	0.000666	0.001845	0.000216	0.000336	0.001423	0.000062	0.000214	0.000497	0.001636	0.000482	0.002886	0.001920	0.000895	0.000324	0.000644	0.000466	0.001437	0.001016	0.000087	0.000028	0.001323	0.000012	0.000294	0.000592	0.000006	0.000000	0.000000	0.000062	0.000000	0.000000
ь	0.000847	0.000556	0.001573	0.000318	0.000519	0.002268	0.000076	0.000484	0.001075	0.000000	0.002123	0.000171	0.000688	0.002501	0.000889	0.001173	0.000265	0.002451	0.000623	0.000254	0.000052	0.000084	0.000123	0.000504	0.000567	0.000011	0.000000	0.000000	0.000210	0.000611	0.000000
э	0.000000	0.000000	0.000003	0.000004	0.000007	0.000000	0.000000	0.000001	0.000000	0.000016	0.000060	0.000106	0.000003	0.000036	0.000000	0.000031	0.000018	0.000056	0.002504	0.000000	0.000000	0.000004	0.000001	0.000001	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
ю	0.000169	0.000605	0.000356	0.000146	0.000706	0.000085																									

2. За допомогою програми CoolPinkProgram оцінити значення (10) Н, (20) Н, (30) Н.

Лабораторная работа №1

Произвольная часть текста:
телом_человек_подел

Использованные буквы:

Порядок n-граммы:
5 символов
10 символов
15 символов
25 символов
30 символов
35 символов
40 символов
45 символов
50 символов

Введенный символ:

Символ по счету:

Номер эксперимента: 50

Неравенство для энтропии:
 $2.79957414891427 < H < 3.33450223668496$

Двоичная таблица угаданных символов:

Поле ввода символов:

Продолжить Другой

Вероятности:

q[1] = 0.3877551
q[2] = 0.1224489
q[3] = 0.0204081
q[4] = 0.0204081
q[5] = 0.0204081
q[6] = 0.0408163
q[7] = 0
q[8] = 0.0204081
q[9] = 0
q[10] = 0
q[11] = 0.020408
q[12] = 0.020408
q[13] = 0
q[14] = 0.040816
q[15] = 0.020408
q[16] = 0
q[17] = 0
q[18] = 0
q[19] = 0.061224
q[20] = 0
q[21] = 0.061224
q[22] = 0
q[23] = 0
q[24] = 0.020408
q[25] = 0.020408
q[26] = 0
q[27] = 0.020408
q[28] = 0
q[29] = 0.040816
q[30] = 0
q[31] = 0.020408
q[32] = 0.020408

Строка состояния:

Ентропія імовірнісного ансамблю $\langle Z, P \rangle$ – це величина

Лабораторная работа №1

Произвольная часть текста:
ния_только_нашему_плохому_лов

Использованные буквы:

Порядок n-граммы:
5 символов
10 символов
15 символов
20 символов
25 символов
35 символов
40 символов
45 символов
50 символов

Введенный символ:

Символ по счету:

Номер эксперимента: 50

Неравенство для энтропии:
 $3.61662761593027 < H < 3.97245996018797$

Двоичная таблица угаданных символов:

Поле ввода символов:

Продолжить Другой

Вероятности:

q[1] = 0.2653061
q[2] = 0.0408163
q[3] = 0.0204081
q[4] = 0.0408163
q[5] = 0.1020408
q[6] = 0.0204081
q[7] = 0
q[8] = 0.0204081
q[9] = 0
q[10] = 0
q[11] = 0.020408
q[12] = 0.040816
q[13] = 0
q[14] = 0.040816
q[15] = 0.040816
q[16] = 0.020408
q[17] = 0.020408
q[18] = 0
q[19] = 0
q[20] = 0.061224
q[21] = 0
q[22] = 0
q[23] = 0.040816
q[24] = 0.020408
q[25] = 0.020408
q[26] = 0.040816
q[27] = 0.020408
q[28] = 0.020408
q[29] = 0.040816
q[30] = 0.020408
q[31] = 0
q[32] = 0.020408

Строка состояния:

Ентропія імовірнісного ансамблю $\langle Z, P \rangle$ – це величина

Лабораторная работа №1

Произвольная часть текста:
как_глубоко_нравится_нам_это_или_нет_верим_мы_в_закон_человеческой_природы_

Использованные буквы:

Порядок n-граммы:
5 символов
15 символов
20 символов
25 символов
30 символов
35 символов
40 символов
45 символов
50 символов

Введенный символ: к

Символ по счету: 1

Номер эксперимента: 50

Неравенство для энтропии:
 $1.90050494509215 < H < 2.64602143851991$

Двоичная таблица угаданных символов:

Поле ввода символов:

Продолжить Другой

Вероятности:

q[1] = 0.54
q[2] = 0.08
q[3] = 0
q[4] = 0.06
q[5] = 0.06
q[6] = 0.04
q[7] = 0.02
q[8] = 0
q[9] = 0.04
q[10] = 0.02
q[11] = 0
q[12] = 0
q[13] = 0.02
q[14] = 0
q[15] = 0
q[16] = 0.02
q[17] = 0
q[18] = 0
q[19] = 0.02
q[20] = 0
q[21] = 0
q[22] = 0
q[23] = 0.02
q[24] = 0
q[25] = 0.02
q[26] = 0
q[27] = 0
q[28] = 0.02
q[29] = 0
q[30] = 0
q[31] = 0
q[32] = 0.02

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

Ентропія імовірнісного ансамблю $\langle Z, P \rangle$ – це величина

3. Використовуючи отримані значення ентропії, оцінити надлишковість російської мови в різних моделях джерела.

$$R = 1 - \frac{H_{\infty}}{H_0} \quad H_0 = \log_2 32 = 5$$

Без пробіла, H_1 :

$$R = 0.108884$$

З пробілом, H_1 :

$$R = 0.21162$$

Перетинаються, з пробілом, H_1 :

$$R = 0.215152$$

Не перетинаються, з пробілом, H_2 :

$$R = 0.230467$$

Перетинаються, без пробіла, H_1 :

$$R = 0.177885$$

Не перетинаються, без пробіла, H_2 :

$$R = 0.17854$$

Код програми

```
#include <iostream>
#include <fstream>
#include <string>
#include <algorithm>
#include <iomanip>
#include <cmath>

using namespace std;
double sum = 0;
double entropy(double tmp)
{
    if (tmp == 0.) return 0;
    return -tmp * log2( tmp );
}

void quickSort(int arr[], char strLC[], int left, int right) {

    int i = left, j = right;

    int tmp;
    char ctmp;

    int pivot = arr[(left + right) / 2];
    while (i <= j) {

        while (arr[i] < pivot)

            i++;

        while (arr[j] > pivot)

            j--;

        if (i <= j) {

            ctmp = strLC[i];
            tmp = arr[i];

            strLC[i] = strLC[j];
            arr[i] = arr[j];

            strLC[j] = ctmp;
            arr[j] = tmp;

            i++;

            j--;

        }

    };
    if (left < j)

        quickSort(arr, strLC, left, j);

    if (i < right)

        quickSort(arr, strLC, i, right);

}
```

```

void copytext()
{
    ofstream output("output.txt");
    ifstream input("input.txt");

    string str, str1 = "";
    bool flag = 0;
    const int n = 34;
    int length;
    char strLC[n] = { ' ', 'а', 'б', 'в', 'г', 'д', 'е', 'ё', 'ж', 'з', 'и', 'й', 'к',
'л', 'м', 'н', 'о', 'п', 'р', 'с', 'т', 'у', 'ф', 'х', 'ц', 'ч', 'ш', 'щ', 'ъ', 'ы',
'ь', 'э', 'ю', 'я' };
    char strUC[n] = { ' ', 'А', 'Б', 'В', 'Г', 'Д', 'Е', 'Ё', 'Ж', 'З', 'И', 'Й', 'К',
'Л', 'М', 'Н', 'О', 'П', 'Р', 'С', 'Т', 'У', 'Ф', 'Х', 'Ц', 'Ч', 'Ш', 'Щ', 'Ъ', 'Ы',
'Ь', 'Э', 'Ю', 'Я' };

    while (input)
    {
        getline(input, str);
        length = str.length();
        for (int i = 0; i < length; i++)
        {
            flag = 0;
            for (int j = 0; j < n; j++)
                if (str[i] == strLC[j] || str[i] == strUC[j])
                {
                    if (j == 7) { str1 += strLC[j - 1]; flag = 1; break;}
                    if(j == 28){ str1+=strLC[30]; flag = 1; break;}
                    str1 += strLC[j];
                    flag = 1;
                    break;
                }

            if (!flag) { str.erase(i, 1); i--; length--; }

        }
        str1 += " ";
        if (input.eof())break;
    }

    string strtemp = "";
    length = str1.length() - 1;
    for (int i = 0; i < length; i++)
    {
        strtemp = str1[i];
        strtemp += str1[i + 1];
        if (strtemp == " ") { str1.erase(i, 1); i--; length--; }
    }

    output << str1 << endl;

    input.close();
    output.close();
}

```

```

void monog()
{
    ofstream output("monog.txt");
    ifstream input("output.txt");

    string str;

    const int n = 32;

```

```

    int length;
    double lengfull = 0.0, tmp = 0.0;
    char strLC[n] = { ' ', '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', 'ы', 'ь', 'э',
'ю', 'я' };

    int arr[n];
    for (int i = 0; i < n; i++)
        arr[i] = 0;
    while (input)
    {
        getline(input, str);
        length = str.length();
        lengfull += length;
        for (int i = 0; i < length; i++)
        {
            for (int j = 0; j < n; j++)
                if (str[i] == strLC[j]) arr[j]++;

        }

    }

    int snum = arr[0]; // number of spaces in text

    quickSort(&arr[0], &strLC[0], 0, n - 1);

    output <<setw(6) << "With space: " <<" " << endl;
    for (int j = n-1; j >= 0; j--)
    {
        if (strLC[j] == ' ')continue;
        tmp = arr[j] / lengfull;
        sum = sum + entropy(tmp);
        output << strLC[j] << " - " << std::setprecision(6) << tmp <<setw(6) << " " << endl;
    }
    output << "Entropia - " << sum << setw(6) << endl;

    output <<setw(6) << "\nWithout space: " <<" " << endl;
    sum = 0;
    for (int j = n-1; j >= 0; j--)
    {
        if (strLC[j] == ' ')continue;
        tmp = arr[j] / (lengfull - snum);
        sum = sum + entropy(tmp);
        output << strLC[j] << " - " << std::setprecision(6) << tmp <<setw(6) << " " << endl;
    }
    output << "Entropia - " << sum << setw(6) << endl;
    input.close();
    output.close();
}

void bigr_ws_wi() // Перетин з пробілами
{
    ofstream output("bigr_ws_wi.txt");
    ifstream input("output.txt");//changed

    string str;
    const int n = 32;
    int length;
    double tmp = 0.0;
    char strLC[n] = { ' ', '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', 'ы', 'ь', 'э',
'ю', 'я' };
    int arr[n * n];

    for (int i = 0; i < n * n; i++)
        arr[i] = 0;

```

```

double count = 0.0;

string arr1[n * n];
for (int i = 0; i < n * n; i++)
    arr1[i] = " "; // Типу весь массив забив нулями
(п о д в і й н и м и п р о б і л а м и)

while (input)
{
    getline(input, str);
    length = str.length();

    for (int i = 0; i < length; i++)
    {
        if (str[i] == ' ' || str[i] == EOF) continue;

        for (int k = 0; k < n; k++)
            if (str[i] == strLC[k])
            {
                if (str[i + 1] == ' ' || str[i + 1] == EOF) { i++; continue; }
                for (int j = 0; j < n; j++)
                    if (str[i + 1] == strLC[j])
                    {
                        arr[k * n + j] += 1;
                        count++;
                        break;
                    }
                break;
            }
    }
    if (input.eof()) break;
}

output << fixed << " "; // Matrix for frequency
for (int i = 1; i < 32; i++)
    output << setw(9) << strLC[i] << " ";
output << "\n\n\n";
for (int i = 1; i < 32; i++)
{
    output << strLC[i] << " ";
    for (int j = 1; j < 32; j++)
        output << setw(9) << floor((arr[i * n + j] / count) * 1000000.) / 1000000. << "
";

    output << "\n\n\n";
}

output << "\n\n\n\n\n";
sum = 0;
for (int i = 1; i < 32; i++) // Matrix for entropy
    output << setw(9) << strLC[i] << " ";
output << "\n\n\n";
for (int i = 1; i < 32; i++)
{
    output << strLC[i] << " ";
    for (int j = 1; j < 32; j++)
    {
        tmp = floor((arr[i * n + j] / count) * 1000000.) / 1000000.;
        sum = sum + entropy(tmp);
        output << setw(9) << entropy(tmp) << " ";
    }
    output << "\n\n\n";
}

```

```

        output <<setw(9) << "Entropy = " << sum << endl;

        input.close();
        output.close();
    }

void bigr_ws_ni() // Без перетину з пробілами
{
    ofstream output("bigr_ws_ni.txt");
    ifstream input("output.txt");

    string str;
    const int n = 32;
    int length;
    double tmp = 0.0;
    char strLC[n] = { ' ', 'a', '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', 'ы', 'ь', 'э',
        'ю', 'я' };
    int arr[n * n];

    for (int i = 0; i < n * n; i++)
        arr[i] = 0;

    double count = 0.0;

    string arr1[n * n];
    for (int i = 0; i < n * n; i++)
        arr1[i] = " ";

    while (input)
    {
        getline(input, str);
        length = str.length();

        for (int i = 0; i < length; i = i + 2)
        {
            if (str[i] == ' ' || str[i] == EOF) continue;

            for (int k = 0; k < n; k++)
                if (str[i] == strLC[k])
                {
                    if (str[i + 1] == ' ' || str[i + 1] == EOF) { i++; continue; }
                    for (int j = 0; j < n; j++)
                        if (str[i + 1] == strLC[j])
                        {
                            if (arr[k * n + j] == 0) arr1[k * n + j] = strLC[k]
+ strLC[j];

                            arr[k * n + j] += 1;
                            count++;
                            break;
                        }
                    break;
                }

            }

        }

        if (input.eof()) break;

    }

    output << fixed << " ";
    for (int i = 1; i < 32; i++) //Matrix for frequency

```



```

        output << setw(9) << strLC[i] << " ";
output << "\n\n\n";
for (int i = 1; i < 32; i++)
{
    output << strLC[i] << " ";
    for (int j = 1; j < 32; j++)
        output << setw(9) << floor((arr[i * n + j] / count) * 1000000.) / 1000000. << "
";

    output << "\n\n\n";
}

output << "\n\n\n\n\n\n";
sum = 0;
for (int i = 1; i < 32; i++) //Matrix for entropy
    output << setw(9) << strLC[i] << " ";
output << "\n\n\n";
for (int i = 1; i < 32; i++)
{
    output << strLC[i] << " ";
    for (int j = 1; j < 32; j++)
    {
        tmp = floor((arr[i * n + j] / count) * 1000000.) / 1000000.;
        sum = sum + entropy(tmp);
        output << setw(9) << entropy(tmp) << " ";
    }
    output << "\n\n\n";
}
output << setw(9) << "Entropy = " << sum << endl;

input.close();
output.close();
}

```

void bigr_ns_wi() // П е р е т и н б е з п р о б і л і в

```

{
    ofstream output("bigr_ns_wi.txt");
    ifstream input("output_ns.txt");
    double tmp = 0.0;
    string str;
    const int n = 32;
    int    length;

    char strLC[n] = { ' ', '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', 'ы', 'ь', 'э',
'ю', 'я' };
    int arr[n * n];

    for (int i = 0; i < n * n; i++)
        arr[i] = 0;

    double count = 0.0;

    while (input)
    {
        getline(input, str);
        length = str.length();

        for (int i = 0; i < length; i++)
        {
            if (str[i] == EOF)continue;

            for (int k = 1; k < n; k++)
                if (str[i] == strLC[k])
                {
                    if (str[i + 1] == EOF) { i++; continue; }
                    for (int j = 1; j < n; j++)

```

```

        if (str[i + 1] == strLC[j])
        {
            arr[k * n + j] += 1;
            count++;
            break;
        }
        break;
    }

    if (input.eof())break;

}

output << fixed << " ";
for (int i = 1; i < 32; i++) //For Freq
    output << setw(9) << strLC[i] << " ";
output << "\n\n";
for (int i = 1; i < 32; i++)
{
    output << strLC[i] << " ";
    for (int j = 1; j < 32; j++)
        output << setw(9) << floor((arr[i * n + j] / count) * 1000000.) / 1000000. << "
";
    output << "\n\n";
}

output << "\n\n\n\n";
sum = 0;
for (int i = 1; i < 32; i++) //Matrix for entropy
    output << setw(9) << strLC[i] << " ";
output << "\n\n";
for (int i = 1; i < 32; i++)
{
    output << strLC[i] << " ";
    for (int j = 1; j < 32; j++)
    {
        tmp = floor((arr[i * n + j] / count) * 1000000.) / 1000000.;
        sum = sum + entropy(tmp);
        output << setw(9) << entropy(tmp) << " ";
    }
    output << "\n\n";
}
output << setw(9) << "Entropy = " << sum << endl;
input.close();
output.close();
}

```

```

void bigr_ns_ni() // Без перетину без пробілів
{

```

```

    ofstream output("bigr_ns_ni.txt");
    ifstream input("output_ns.txt");

    string str;
    const int n = 32;
    int length;
    double tmp = 0.0;
    char strLC[n] = { ' ', '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', 'ы', 'ь', 'э',
'ю', 'я' };
    int arr[n * n];

    for (int i = 0; i < n * n; i++)

```

```

        arr[i] = 0;

double count = 0.0;

while (input)
{
    getline(input, str);
    length = str.length();

    for (int i = 0; i < length; i = i + 2)
    {
        if (str[i] == EOF)continue;

        for (int k = 1; k < n; k++)
            if (str[i] == strLC[k])
            {
                if (str[i + 1] == EOF) { i++; continue; }
                for (int j = 1; j < n; j++)
                    if (str[i + 1] == strLC[j])
                    {
                        arr[k * n + j] += 1;
                        count++;
                        break;
                    }
                break;
            }
    }

    if (input.eof())break;
}

output << fixed << " ";
for (int i = 1; i < 32; i++) //For freq
    output << setw(9) << strLC[i] << " ";
output << "\n\n\n";
for (int i = 1; i < 32; i++)
{
    output << strLC[i] << " ";
    for (int j = 1; j < 32; j++)
        output << setw(9) << floor((arr[i * n + j] / count) * 1000000.) / 1000000. << "
";

    output << "\n\n\n";
}

output << "\n\n\n\n\n\n";
sum = 0;
for (int i = 1; i < 32; i++) //Matrix for entropy
    output << setw(9) << strLC[i] << " ";
output << "\n\n\n";
for (int i = 1; i < 32; i++)
{
    output << strLC[i] << " ";
    for (int j = 1; j < 32; j++)
    {
        tmp = floor((arr[i * n + j] / count) * 1000000.) / 1000000.;
        sum = sum + entropy(tmp);
        output << setw(9) << entropy(tmp) << " ";
    }
    output << "\n\n\n";
}
output << setw(9) << "Entropy = " << sum << endl;

input.close();
output.close();

```

```
}

int main() {

    setlocale(LC_ALL, "Russian");

    copytext();
    monog();
    bigr_ws_wi();
    bigr_ws_ni();
    bigr_ns_wi();
    bigr_ns_ni();

    system("pause");
    return 0;
}
```

Висновок:

Отже, в ході практикума ми засвоїли поняття 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