

НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ УКРАЇНИ
«КИЇВСЬКИЙ ПОЛІТЕХНІЧНИЙ ІНСТИТУТ»
ФІЗИКО-ТЕХНІЧНИЙ ІНСТИТУТ

КРИПТОГРАФІЯ
КОМП'ЮТЕРНИЙ ПРАКТИКУМ №1

Експериментальна оцінка ентропії на символ джерела
відкритого текст

Виконали:
студенти групи ФБ-32
Кошикова Дар'я
Сажко Олена

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

Постановка задачі:

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

Хід роботи

Спочатку ми виконали фільтрацію тексту: привели всі символи до нижнього регістру, замінили непотрібні символи на пробіли та зробили дві версії тексту — з пробілами і без них:

```
def clean_text(text, keep_spaces=True):
    text = text.lower()
    text = re.sub(r'^a-яё\s', ' ', text)
    text = re.sub(r'\s+', ' ', text).strip()
    if not keep_spaces:
        text = text.replace(' ', '')
    return text
```

Також для збереження результатів попередньої обробки ми записали текст у два файли: з пробілами та без пробілів:

```
def save_filtered(raw_text, path_with_spaces, path_no_spaces):
    filtered = clean_text(raw_text, keep_spaces=True)
    filtered_no_spaces = clean_text(raw_text, keep_spaces=False)

    with open(path_with_spaces, "w", encoding="utf-8") as f:
        f.write(filtered)
    with open(path_no_spaces, "w", encoding="utf-8") as f:
        f.write(filtered_no_spaces)

    return filtered, filtered_no_spaces
```

Далі ми обрахували частоти всіх літер у тексті:

```
def letter_freq(text):
    counts = Counter(text)
    total = len(text)
    freqs = {c: count / total for c, count in counts.items()}
    return dict(sorted(freqs.items(), key=lambda x: x[1], reverse=True))
```

Для аналізу залежностей ми підраховували частоти біграм. Розглядали біграми, що перетинаються (step=1), і ті, що не перетинаються (step=2):

```
def bigram_freq(text, step=1):
    bigrams = [text[i:i+2] for i in range(0, len(text)-1, step) if
len(text[i:i+2]) == 2]
    counts = Counter(bigrams)
    total = len(bigrams)
    freqs = {bg: count / total for bg, count in counts.items()}
    return freqs
```

Частота літер в тексті (з пробілами):

Літера	Частота
	0.176762
о	0.089328
и	0.075581
е	0.073173
а	0.061429
с	0.047335
т	0.047247
н	0.045687
в	0.041796
л	0.036486
р	0.035425
д	0.02942
м	0.027242
п	0.022289
к	0.021807
у	0.0213
г	0.018223
я	0.016809
ы	0.015779
б	0.015313
з	0.013409
ь	0.012393
х	0.009437
ч	0.008764
й	0.007751
ж	0.007363
ш	0.006183
ю	0.005786
ц	0.004518
щ	0.002928
ф	0.001994
э	0.000841
ъ	0.000114
ё	8.54E-05

Частота літер в тексті (без пробілів):

Літера	Частота
о	0.10851
и	0.09181
е	0.08888
а	0.07462
с	0.0575
т	0.05739
н	0.0555
в	0.05077
л	0.04432
р	0.04303
д	0.03574
м	0.03309
п	0.02707
к	0.02649
у	0.02587
г	0.02214
я	0.02042
ы	0.01917
б	0.0186
з	0.01629
ь	0.01505
х	0.01146
ч	0.01065
й	0.00942
ж	0.00894
ш	0.00751
ю	0.00703
ц	0.00549
щ	0.00356
ф	0.00242
э	0.00102
ъ	0.00014
ё	0.0001

Перетинаючі біграми:

	а	б	в	г	д	е	ё	ж	з	и	й	к	л	м	н	о	п	р
а	0.000397	0.000626	0.004143	0.000956	0.001981	0.001185	3.77E-06	0.000868	0.002836	0.000898	0.000777	0.003319	0.006168	0.00356	0.003075	0.000166	0.000632	0.004138
б	0.000442	5.52E-05	4.88E-05	1.42E-06	1.04E-05	0.001859	0	2.36E-06	1.18E-06	0.000717	0	1.27E-05	0.001244	2.71E-05	0.000279	0.003473	0	0.001369
в	0.005045	3.07E-06	9.81E-05	4.72E-06	0.00023	0.005527	5.19E-06	0	0.000387	0.003055	0	3.99E-05	0.000785	0.000135	0.00053	0.009392	6.98E-05	0.001032
г	0.001103	1.89E-06	6.37E-06	1.63E-05	0.001093	0.00031	0	0	0	0.000784	0	9.91E-06	0.000875	9.44E-07	0.000473	0.011541	2.36E-07	0.000704
д	0.005323	7.31E-06	0.000618	1.42E-06	4.39E-05	0.006125	3.77E-06	2.36E-07	1.37E-05	0.002549	0	6.11E-05	0.000799	5.66E-05	0.001855	0.003441	2.85E-05	0.000818
е	3.77E-05	0.002196	0.002282	0.004578	0.00271	0.00122	2.83E-06	0.000855	0.001198	0.000208	0.00231	0.001102	0.00572	0.005251	0.006721	9.77E-05	0.00064	0.005008
ё	0	0	0	2.36E-07	4.72E-07	0	0	0	2.12E-06	0	0	4.72E-07	5.66E-06	5.43E-06	4.95E-06	0	0	2.59E-06
ж	0.00064	3.35E-05	2.59E-06	2.17E-05	0.001017	0.002938	7.08E-07	9.95E-05	0	0.001804	0	2.15E-05	1.18E-06	1.89E-06	0.000396	1.72E-05	0	5.31E-05
з	0.003965	0.000247	0.000872	0.000153	0.000658	0.001079	0	1.25E-05	0.000138	0.000248	0	3.66E-05	0.000482	0.000163	0.001064	0.000527	0	0.000942
и	0.00038	0.001121	0.001579	0.000301	0.001686	0.003441	0	0.00019	0.002903	0.00124	0.001132	0.001924	0.005596	0.003905	0.002727	0.000774	0.000253	0.000968
й	0	0	7.08E-07	2.36E-07	0.000415	4.72E-07	0	0	2.36E-07	1.18E-06	0	3.77E-06	1.18E-06	1.75E-05	0.00019	7.08E-07	0	0
к	0.004958	4.01E-06	0.000103	7.08E-07	2.36E-07	0.000261	0	8.96E-05	0	0.001898	0	7.31E-05	0.000543	2.36E-06	0.000233	0.006422	0	0.001445
л	0.004976	3.66E-05	1.27E-05	6.25E-05	3.44E-05	0.004404	1.65E-06	0.000267	2.83E-06	0.007542	2.36E-07	0.00036	8E-05	4.29E-05	0.000357	0.004533	4.86E-05	0
м	0.001678	2.36E-07	1.53E-05	7.55E-06	7.08E-07	0.003654	4.72E-07	0	2.59E-06	0.002974	0	0.000191	0.001021	6.68E-05	0.00139	0.003233	1.27E-05	9.77E-05
н	0.009626	0	1.42E-06	0.000139	2.81E-05	0.008136	2.59E-06	2.36E-07	6.37E-06	0.008745	0	6.39E-05	7.08E-07	6.37E-06	0.001877	0.005967	0	1.11E-05
о	0.000238	0.003056	0.007237	0.004947	0.007044	0.002844	3.07E-06	0.001539	0.001334	0.002093	0.002184	0.001685	0.003718	0.003858	0.004825	5.97E-05	0.000763	0.006015
п	0.001311	0	0	0	0	0.001377	7.08E-07	0	0	0.00081	0	4.27E-05	0.000746	0	6.65E-05	0.009379	2.76E-05	0.00631
р	0.006887	9.03E-05	0.000232	0.000148	0.000523	0.005926	1.89E-06	0.000187	9.77E-05	0	0.00478	0	7.34E-05	5.76E-05	0.000406	0.00618	8.68E-05	5.43E-05
с	0.001924	4.69E-05	0.002591	1.18E-05	0.000499	0.004545	5.12E-05	1.84E-05	4.48E-06	0.001868	2.36E-07	0.003435	0.003203	0.000484	0.000646	0.002387	0.003202	0.000296
т	0.004071	1.42E-05	0.003857	3.77E-06	0.000168	0.004733	9.44E-07	2.36E-07	4.72E-07	0.003141	0	0.000414	8.99E-05	3.49E-05	0.000481	0.009783	0.000137	0.002039
у	1.06E-05	0.000498	0.000388	0.000686	0.003058	0.000168	0	0.000872	0.000211	8.4E-05	0.000109	0.000772	0.003032	0.00076	0.000116	1.18E-06	0.000673	0.000184
ф	0.00049	0	1.18E-06	2.36E-07	4.72E-07	0.000193	0	0	0	0.000233	0	1.65E-06	4.51E-05	1.89E-06	7.08E-06	9.6E-05	0	7.19E-05
х	0.000524	3.07E-06	0.000133	4.72E-07	7.08E-07	0.000179	0	0	1.65E-06	0.000275	2.36E-07	2.36E-07	0.00018	1.49E-05	3.56E-05	0.001316	2.12E-06	0.000462
ц	0.002736	4.72E-07	3.35E-05	3.07E-06	0	0.000785	0	0	0	2.78E-05	0	1.89E-06	7.08E-07	7.08E-07	0	0.000152	1.42E-06	3.3E-06
ч	0.001128	2.36E-07	1.89E-06	0	0	0.002285	7.08E-07	0	0	0.001407	0	3E-05	1.72E-05	9.67E-06	0.000374	4.41E-05	0	0.000122
ш	0.000857	4.48E-06	5.66E-06	2.36E-07	0	0.001773	5.19E-06	0	0	0.00162	0	3.02E-05	0.000529	2.83E-06	0.000226	9.81E-05	4.72E-07	7.08E-07
щ	0.000429	0	1.82E-05	0	0	0.00125	7.08E-07	0	0	0.000992	0	0	0	0	4.69E-05	7.08E-07	0	5.43E-06
ъ	0	0	0	0	0	3.35E-05	0	0	0	9.44E-07	3.07E-06	0	0	0	0	0	0	0
ы	0	4.55E-05	0.000645	1.72E-05	8.42E-05	0.001048	0	1.84E-05	2.36E-06	0.001148	0.000147	0.001272	0.000751	0.001647	0	4.34E-05	0.000164	0
ь	0	1.39E-05	2.76E-05	1.11E-05	9.11E-05	0.000196	0	4.74E-05	2.48E-05	0	0.000212	0	0.000351	0.000557	0	4.72E-07	0	0
э	0	2.36E-07	4.72E-07	1.42E-06	4.72E-07	0	0	0	0	4.72E-07	2.36E-07	6.84E-06	1.18E-06	9.44E-07	0	9.44E-07	3.3E-06	0
ю	0	0.00024	0	2.17E-05	0.000334	1.18E-06	0	1.75E-05	2.22E-05	0	9.44E-07	2.12E-06	3.54E-06	1.42E-06	5.83E-05	0	4.72E-07	3.07E-06
я	1.18E-06	1.67E-05	0.000232	3.96E-05	0.000102	0.000184	0	6.25E-05	0.000302	7.08E-06	8.52E-05	0.000279	0.000593	0.0005	0.00084	0	4.72E-07	5.38E-05
	0.002253	0.006951	0.016608	0.006088	0.007608	0.005513	0	0.002262	0.003882	0.02555	1.18E-06	0.007461	0.002402	0.007159	0.014182	0.010247	0.015665	0.003049

р	с	т	у	ф	х	ц	ч	ш	щ	ъ	ы	ь	э	ю	я	
0.004138	0.002788	0.002846	0.000167	0.000312	0.001103	2.24E-05	0.000599	0.001052	0.000219	0	0	0	7.08E-07	0.000896	0.000651	0.015044
0.001369	4.29E-05	4.01E-06	0.001606	0	8.49E-06	3.3E-06	9.44E-07	8.73E-06	6.86E-05	7.1E-05	0.00293	4.13E-05	0	0	0.000726	0.000259
0.001032	0.002567	0.000468	0.000875	9.91E-06	7.74E-05	6.35E-05	7.27E-05	0.000531	1.42E-06	0	0.002283	0.000359	0	2.36E-07	0.000847	0.007303
0.000704	1.25E-05	1.89E-06	0.000433	4.72E-07	0	0	1.13E-05	1.18E-05	0	0	0	0	0	0	0	0.000833
0.000818	0.000237	7.17E-05	0.002225	0	9.2E-06	0.00056	6.84E-06	7.22E-05	1.49E-05	4.48E-06	0.000755	0.001145	0	2.59E-06	0.000451	0.002122
0.005008	0.004744	0.005063	6.77E-05	0.000326	0.000783	0.000327	0.001031	0.00083	0.000346	0	0	0	2.12E-06	0.000539	0.000493	0.016483
2.59E-06	0	5.66E-06	0	0	1.18E-06	0	0	2.36E-07	2.36E-07	0	0	0	0	4.72E-07	0	5.57E-05
5.31E-05	3.54E-06	0	0.000194	0	0	0	1.79E-05	0	0	0	0	3.66E-05	0	0	0	6.39E-05
0.000942	1.42E-06	2.59E-06	0.000239	2.36E-07	0	2.59E-06	1.18E-06	4.72E-07	0	8.26E-06	0.00039	0.000185	2.36E-07	8.73E-06	0.000288	0.001696
0.000968	0.004487	0.004422	0.000457	0.000184	0.003304	0.000705	0.00037	0.000568	0.00043	0	0	0	1.18E-06	0.000378	0.001711	0.028443
0	0.000545	0.000431	0	0	0	3.09E-05	1.18E-06	9.32E-05	0	0	0	0	0	2.36E-07	6.13E-06	0.006012
0.001445	6.42E-05	0.000532	0.000746	4.72E-07	1.32E-05	0	0	6.37E-06	0	0	0	0	0	0	0	0.004413
0	0.000628	2.38E-05	0.001152	4.72E-06	2.5E-05	9.44E-07	6.93E-05	6.37E-06	1.42E-06	0	0.000532	0.00174	0	0.001059	0.001737	0.006744
9.77E-05	0.000181	7.08E-07	0.002606	0.000381	2.36E-07	8.49E-06	5.66E-06	1.18E-06	1.84E-05	1.89E-06	0.000728	0.000175	0	0	0.00056	0.008232
1.11E-05	0.000213	6.65E-05	0.001107	4.72E-06	6.13E-06	0.000215	7.43E-05	2.36E-07	8.82E-05	2.36E-07	0.002736	0.000832	7.08E-07	0.000125	0.001513	0.004093
0.006015	0.007488	0.007468	6.98E-05	0.000203	0.000334	9.44E-05	0.000941	0.000758	0.000157	0	0	0	2.12E-05	0.001224	0.000493	0.016632
0.00631	0.000698	0.00016	0.000752	6.6E-06	0	1.16E-05	7.08E-06	2.57E-05	2.59E-06	0	0.000129	6.23E-05	0	0	0.000302	6.04E-05
5.43E-05	0.00044	0.000645	0.0002136	1.42E-05	8.56E-05	3.89E-05	2.45E-05	0.000215	1.91E-05	7.08E-07	0.001567	0.000523	0	0.000221	0.000814	0.002879
0.000296	0.000538	0.009138	0.001094	6.09E-05	0.000573	6.56E-05	0.000145	4.53E-05	2.36E-07	2.62E-05	0.001507	0.001678	0	8.42E-05	0.003263	0.003904
0.002039	0.00126	9.06E-05	0.000787	8.96E-06	1.77E-05	0.000326	0.000331	7.78E-06	1.91E-05	9.44E-07	0.001963	0.004653	0	1.65E-06	0.000487	0.008325
0.000184	0.001786	0.001315	0	3.56E-05	0.000449	1.53E-05	0.000569	0.000667	0.000294	0	0	0	1.65E-06	0.00067	3.56E-05	0.006569
7.19E-05	1.98E-05	3.77E-06	5.17E-05	2.59E-05	7.08E-07	7.08E-07	0	2.36E-07	0	0	2.15E-05	4.76E-05	0	0	8.73E-06	0.000672
0.000462	6.06E-05	2.36E-07	0.000164	4.72E-07	0	9.44E-07	0	2.12E-06	0	0	0	2.36E-07	0	2.36E-07	0	0.006082
3.3E-06	0	0	0.000162	1.89E-06	0	7.08E-07	0	0	0	0	0.000237	0	0	0	0	0.00037
0.000122	0	0.002396	0.000232	0	0	2.12E-06	0	6.98E-05	0	0	0	0.000218	0	0</		

Непертинаючі біграми:

	а	б	в	г	д	е	ё	ж	з	и	й	к	л	м	н	о	п	р
а	0.000414	0.000622	0.004116	0.000962	0.001973	0.001196	3.77E-06	0.000874	0.00285	0.000892	0.00077	0.003247	0.006227	0.003516	0.003079	0.000165	0.000642	0.004093
б	0.000437	5.9E-05	5E-05	1.89E-06	1.09E-05	0.001844	0	2.36E-06	1.42E-06	0.000724	0	1.18E-05	0.001223	2.88E-05	0.000273	0.003488	0	0.001375
в	0.005087	2.83E-06	0.000103	5.66E-06	0.000251	0.005485	6.13E-06	0	0.000413	0.003056	0	3.87E-05	0.0008	0.000136	0.000536	0.009497	6.89E-05	0.001033
г	0.001095	2.36E-06	4.25E-06	1.93E-05	0.0011	0.00033	0	0	0	0.00079	0	1.13E-05	0.000881	9.44E-07	0.000468	0.011477	4.72E-07	0.000707
д	0.005297	5.19E-06	0.000612	0	4.58E-05	0.006132	4.72E-06	4.72E-07	1.42E-05	0.002555	0	6.23E-05	0.000808	5.9E-05	0.001855	0.003402	2.59E-05	0.000835
е	4.2E-05	0.00218	0.002299	0.004594	0.002735	0.001227	3.3E-06	0.000882	0.001222	0.000213	0.002326	0.001097	0.005708	0.005224	0.006692	0.0001	0.000651	0.004986
ё	0	0	0	4.72E-07	4.72E-07	0	0	0	2.83E-06	0	0	0	9.44E-07	6.6E-06	6.13E-06	3.77E-06	0	1.89E-06
ж	0.000632	3.16E-05	3.77E-06	2.36E-05	0.001018	0.002983	0	0.000107	0	0.001771	0	2.12E-05	1.42E-06	1.42E-06	0.000393	2.03E-05	0	4.81E-05
з	0.003924	0.00025	0.000881	0.000151	0.00066	0.001083	0	1.32E-05	0.000129	0.000242	0	3.3E-05	0.000485	0.000157	0.001049	0.000536	0	0.000943
и	0.000381	0.00111	0.00159	0.000303	0.00168	0.003426	0	0.000197	0.002873	0.001236	0.001118	0.0001935	0.005587	0.003911	0.002724	0.000773	0.000242	0.000982
й	0	0	4.72E-07	0	0.00041	9.44E-07	0	0	4.72E-07	1.42E-06	0	3.3E-06	1.42E-06	2.17E-05	0.000197	4.72E-07	0	0
к	0.004938	4.25E-06	0.000101	9.44E-07	4.72E-07	0.000247	0	9.86E-05	0	0.001883	0	7.36E-05	0.000567	3.3E-06	0.000235	0.006467	0	0.001426
л	0.004982	3.73E-05	1.51E-05	6.18E-05	3.49E-05	0.004384	1.89E-06	0.000255	3.3E-06	0.007521	4.72E-07	0.000396	7.22E-05	4.1E-05	0.000353	0.004488	4.62E-05	0
м	0.001699	4.72E-07	1.6E-05	8.49E-06	9.44E-07	0.00362	0	0	2.36E-06	0.003013	0	0.000193	0.001036	6.7E-05	0.001409	0.003243	1.18E-05	9.01E-05
н	0.009683	0	9.44E-07	0.000129	2.78E-05	0.008089	3.3E-06	0	6.6E-06	0.008799	0	6.56E-05	4.72E-07	5.66E-06	0.001912	0.005933	0	1.04E-05
о	0.000235	0.003104	0.0007206	0.004972	0.006976	0.002859	3.3E-06	0.001522	0.001345	0.002122	0.002186	0.00168	0.003722	0.003866	0.004788	5.94E-05	0.00076	0.006012
п	0.001327	0	0	0	0	0.001401	9.44E-07	0	0	0.000817	0	4.25E-05	0.000753	0	6.89E-05	0.009306	2.97E-05	0.00635
р	0.006904	9.01E-05	0.00023	0.000145	0.000541	0.005899	1.89E-06	0.000186	0.0001	0.004737	0	7.6E-05	5.28E-05	7.27E-05	0.000404	0.006174	8.92E-05	5E-05
с	0.001905	4.81E-05	0.002584	1.13E-05	0.000505	0.004577	5.14E-05	1.84E-05	4.72E-06	0.001858	0	0.003467	0.00319	0.000484	0.000647	0.002444	0.003187	0.000296
т	0.004142	1.6E-05	0.003801	5.66E-06	0.000174	0.00469	9.44E-07	0	9.44E-07	0.003098	0	0.000426	8.92E-05	3.77E-05	0.000467	0.009719	0.000131	0.002078
у	9.91E-06	0.000498	0.000388	0.000697	0.00307	0.000173	0	0.000865	0.000219	7.64E-05	0.000111	0.000756	0.000308	0.000756	0.000109	9.44E-07	0.000685	0.000185
ф	0.000483	0	1.42E-06	4.72E-07	0	0.000182	0	0	0	0.000239	0	1.89E-06	4.25E-05	1.89E-06	6.13E-06	9.01E-05	0	7.45E-05
х	0.000524	1.89E-06	0.000126	9.44E-07	1.42E-06	0.000165	0	0	2.36E-06	0.000272	0	4.72E-07	0.000188	1.46E-05	3.59E-05	0.001303	1.42E-06	0.000461
ц	0.002712	4.72E-07	3.26E-05	2.83E-06	0	0.000812	0	0	0	2.41E-05	0	1.89E-06	9.44E-07	4.72E-07	0	0.00015	1.89E-06	2.83E-06
ч	0.001133	0	1.89E-06	0	0	0.002284	4.72E-07	0	0	0.001395	0	3.11E-05	1.79E-05	1.13E-05	0.000388	5.05E-05	0	0.000128
ш	0.000852	4.72E-06	5.66E-06	0	0	0.001802	3.77E-06	0	0	0.001626	0	3.44E-05	0.000531	2.83E-06	0.000233	9.44E-05	9.44E-07	1.42E-06
щ	0.000435	0	2.12E-05	0	0	0.00123	1.42E-06	0	0	0.001009	0	0	0	4.58E-05	4.72E-07	0	0	0.0008E-06
ъ	0	0	0	0	0	3.35E-05	0	0	0	0	2.36E-06	0	0	0	0	0	0	0
ы	0	4.76E-05	0.000638	1.6E-05	8.11E-05	0.001065	0	1.84E-05	1.32E-05	3.3E-06	0.00112	0.000158	0.001276	0.000734	0.001663	0	3.92E-05	0.00017
ь	0	1.6E-05	3.02E-05	1.23E-05	9.11E-05	0.000195	0	0	4.25E-05	2.22E-05	0	0.000213	0	0.000333	0.000545	0	4.72E-07	0
э	0	0	9.44E-07	9.44E-07	9.44E-07	0	0	0	0	0	4.72E-07	0	7.55E-06	1.42E-06	1.42E-06	0	1.42E-06	2.83E-06
ю	0	0.000243	0	2.45E-05	0.000327	1.42E-06	0	1.37E-05	2.31E-05	0	0	2.36E-06	3.77E-06	2.36E-06	5.57E-05	0	4.72E-07	2.83E-06
я	1.42E-06	1.56E-05	0.000237	4.2E-05	9.25E-05	0.000181	0	6.37E-05	0.000298	5.66E-06	8.63E-05	0.00028	0.000617	0.000512	0.000824	0	4.72E-07	5.05E-05
	0.002248	0.006971	0.016519	0.00611	0.007638	0.005484	0	0.002234	0.0039	0.025511	4.72E-07	0.007412	0.002384	0.007146	0.014216	0.010234	0.015608	0.003032

р	с	т	у	ф	х	ц	ч	ш	щ	ъ	ы	ь	э	ю	я	
0.004093	0.002784	0.00286	0.000179	0.00032	0.001069	2.64E-05	0.000597	0.001034	0.000212	0	0	0	9.44E-07	0.000911	0.000628	0.015072
0.001375	3.87E-05	3.77E-06	0.001589	0	9.44E-06	4.25E-06	9.44E-07	8.02E-06	7.45E-05	6.98E-05	0.002918	4.2E-05	0	0	0.000723	0.000253
0.001033	0.002542	0.00046	0.00087	1.04E-05	7.31E-05	5.94E-05	7.97E-05	0.000542	9.44E-07	0	0.002294	0.000362	0	4.72E-07	0.000857	0.007303
0.000707	1.18E-05	1.42E-06	0.000426	4.72E-07	0	0	1.32E-05	8.02E-06	0	0	0	0	0	0	0	0.000797
0.000835	0.00025	6.7E-05	0.002232	0	8.49E-06	0.000561	6.6E-06	6.56E-05	1.37E-05	5.66E-06	0.000757	0.001159	0	1.89E-06	0.000446	0.002107
0.004986	0.004753	0.005084	7.31E-05	0.000315	0.000783	0.000332	0.001019	0.000844	0.000359	0	0	0	2.36E-06	0.000545	0.000492	0.016483
1.89E-06	0	5.19E-06	0	0	9.44E-07	0	0	0	0	0	0	0	0	9.44E-07	0	5.33E-05
4.81E-05	2.83E-06	0	0.000192	0	0	0	1.79E-05	0	0	0	0	3.73E-05	0	0	0	6.98E-05
0.000943	1.42E-06	2.36E-06	0.000237	4.72E-07	0	4.72E-07	1.89E-06	0	0	9.91E-06	0.00038	0.000193	0	8.02E-06	0.000273	0.001707
0.000982	0.004475	0.00442	0.000453	0.000185	0.003343	0.000709	0.000353	0.000566	0.000429	0	0	0	9.44E-07	0.00037	0.001702	0.028576
0	0.000543	0.000429	0	0	3.16E-05	9.44E-07	8.87E-05	0	0	0	0	0	0	0	7.55E-06	0.006044
0.001426	6.32E-05	0.000532	0.000761	4.72E-07	1.13E-05	0	0	5.19E-06	0	0	0	0	0	0	0	0.004424
0	0.000621	2.55E-05	0.00117	4.72E-06	2.45E-05	9.44E-07	6.27E-05	8.02E-06	1.42E-06	0	0.000533	0.001773	0	0.001066	0.001721	0.006682
9.01E-05	0.00018	9.44E-07	0.002591	0.000378	0	8.02E-06	6.6E-06	1.89E-06	1.46E-05	2.36E-06	0.000715	0.000188	0	0	0.000558	0.008273
1.04E-05	0.000206	6.7E-05	0.001097	5.66E-06	7.55E-06	0.00021	7.55E-05	0	8.73E-05	0	0.002715	0.000812	0	0.000118	0.001486	0.004145
0.006012	0.007528	0.007466	6.65E-05	0.000206	0.000347	9.34E-05	0.000944	0.000742	0.000153	0	0	0	2.22E-05	0.001229	0.000489	0.016735
0.00635	0.00073	0.000162	0.000771	5.19E-06	0	9.44E-06	7.55E-06	2.78E-05	2.83E-06	0	0.000125	6.13E-05	0	0	0.000298	5.85E-05
5E-05	0.000431	0.000658	0.002131	1.75E-05	8.87E-05	4.67E-05	2.22E-05	0.000209	1.56E-05	9.44E-07	0.001595	0.00052	0	0.000232	0.000821	0.002874
0.000296	0.000545	0.009157	0.001106	5.71E-05	0.000571	6.51E-05	0.00015	4.53E-05	4.72E-07	2.59E-05	0.001481	0.001656	0	8.54E-05	0.003223	0.003902
0.002078	0.001294	9.11E-05	0.000775	1.13E-05	1.75E-05	0.000338	0.000319	8.49E-06	1.6E-05	4.72E-07	0.001985	0.00467	0	9.44E-07	0.000485	0.00833
0.000185	0.001776	0.001271	0	3.82E-05	0.00046	1.6E-05	0.000573	0.000661	0.000297	0	0	0	1.89E-06	0.000661	3.26E-05	0.006582
7.45E-05	1.89E-05	4.72E-06	5.61E-05	2.74E-05	9.44E-07	4.72E-07	0	0	0	0	2.08E-05	4.86E-05	0	0	9.44E-06	0.000687
0.000461	6.13E-05	0	0.000164	0	0	9.44E-07	0	1.89E-06	0	0	4.72E-07	0	0	0	0	0.00609
2.83E-06	0	0	0.000165	2.36E-06	0	0	0	0	0	0	0.000229	0	0	0	0	0.000368
0.000128	0	0.002453	0.000243	0	0	3.3E-06	0	7.69E-05	0	0	0	0.000215	0	0	0	0.000441
1.42E-06	4.72E-07	3.3E-06	0.000215	9.44E-07	9.44E-07	2.08E-05	0	2.36E-06	0	0	0	0.000671	0	0	0	0.000121
7.08E-06																

Перетинаючі біграми без пробілу:

	а	б	в	г	д	е	ё	ж	з	и	й	к	л	м	н	о	п	р
а	0.000725	0.00153	0.006533	0.001821	0.003129	0.002011	4.58E-06	0.001319	0.003898	0.00388	0.000944	0.00484	0.007715	0.005134	0.0053	0.001154	0.002173	0.005337
б	0.000543	7.88E-05	7.02E-05	8.88E-06	1.63E-05	0.002266	0	6.3E-06	4.01E-06	0.000932	0	2.09E-05	0.001512	4.7E-05	0.000349	0.004282	8.6E-06	0.001666
в	0.006254	0.000196	0.000649	0.000382	0.000805	0.006978	6.3E-06	0.000149	0.00077	0.004891	5.73E-07	0.00047	0.001061	0.000464	0.001257	0.011786	0.000925	0.001561
г	0.001353	4.18E-05	0.000112	5.64E-05	0.001399	0.000414	0	1.09E-05	3.27E-05	0.001096	0	4.53E-05	0.001069	4.3E-05	0.000669	0.014069	6.73E-05	0.000864
д	0.006505	8.88E-05	0.000979	0.000179	0.000117	0.007501	4.58E-06	3.09E-05	6.3E-05	0.003486	0	0.000183	0.001082	0.000162	0.002549	0.004291	0.000177	0.00103
е	0.000221	0.003746	0.004617	0.006167	0.004292	0.002002	3.44E-06	0.001311	0.001979	0.002931	0.002806	0.002067	0.007179	0.007226	0.009576	0.001362	0.002764	0.006369
ё	0	2.29E-06	5.44E-06	1.43E-06	2.58E-06	1.72E-06	0	5.73E-07	3.44E-06	5.44E-06	0	1.72E-06	8.02E-06	8.6E-06	9.74E-06	2.87E-06	5.16E-06	3.15E-06
ж	0.000778	4.5E-05	7.16E-06	2.72E-05	0.001237	0.003578	8.6E-07	0.000122	2.87E-07	0.002202	0	2.98E-05	2.01E-06	5.16E-06	0.000488	2.64E-05	6.02E-06	6.53E-05
з	0.004848	0.000332	0.001222	0.000268	0.000915	0.001367	0	3.98E-05	0.000277	0.000467	0	0.000168	0.000619	0.000293	0.001512	0.000708	0.000147	0.001216
и	0.000814	0.002746	0.005809	0.001183	0.003458	0.00528	0	0.000706	0.00428	0.004756	0.001376	0.003827	0.007174	0.005949	0.006108	0.003403	0.004236	0.001849
й	0.000111	0.000254	0.000627	0.000245	0.000969	0.000185	0	0.000123	0.000215	0.001242	0	0.000319	7.91E-05	0.000321	0.000741	0.000318	0.00061	0.000139
к	0.006086	0.000236	0.00051	0.000333	0.000198	0.000422	0	0.0002	8.19E-05	0.002991	0	0.000366	0.000795	0.000178	0.000808	0.008135	0.000434	0.001846
л	0.006209	0.000292	0.000769	0.000473	0.000355	0.005999	2.01E-06	0.000385	0.000114	0.0010429	2.87E-07	0.000754	0.000176	0.000519	0.000961	0.006026	0.00048	0.000106
м	0.002205	0.00037	0.000886	0.00037	0.000415	0.00474	5.73E-07	9.77E-05	0.000172	0.005631	0	0.000665	0.00135	0.000467	0.002412	0.004411	0.000819	0.000243
н	0.011786	0.000299	0.00039	0.000312	0.00022	0.010046	3.15E-06	7.77E-05	9.03E-05	0.011441	0	0.000252	4.44E-05	0.000201	0.002638	0.007547	0.000466	0.00011
о	0.000489	0.004567	0.011126	0.006609	0.009503	0.003829	3.72E-06	0.002094	0.002041	0.005161	0.002652	0.002886	0.004752	0.005602	0.007487	0.001413	0.002613	0.007602
п	0.001593	1.15E-06	5.16E-06	3.15E-06	2.58E-06	0.001677	8.6E-07	1.43E-06	2.58E-06	0.000997	0	5.67E-05	0.000908	5.73E-07	8.57E-05	0.0114	4.04E-05	0.007665
р	0.008453	0.000169	0.000464	0.0003	0.00074	0.007334	2.29E-06	0.000242	0.000165	0.006859	2.87E-07	0.000219	0.000111	0.000204	0.000698	0.00767	0.000495	0.000123
с	0.002408	0.00018	0.003516	0.000181	0.000799	0.005634	6.22E-05	8.8E-05	9.11E-05	0.003104	2.87E-07	0.004348	0.003964	0.000796	0.001423	0.003153	0.000437	0.00045
т	0.005042	0.000356	0.005631	0.000629	0.000688	0.006139	1.15E-06	0.000138	0.000223	0.004843	0	0.000902	0.000328	0.000499	0.001407	0.012468	0.00098	0.002701
у	0.000124	0.000858	0.001163	0.001087	0.003986	0.000415	0	0.001154	0.000389	0.001444	0.000132	0.001293	0.000452	0.001264	0.000848	0.000319	0.001328	0.000326
ф	0.000617	2.72E-05	3.61E-05	2.55E-05	2.72E-05	0.000263	0	3.72E-06	1.03E-05	0.000452	0	3.72E-05	0.000127	0.000116	4.87E-05	0.000152	4.84E-05	0.000105
х	0.000739	0.00025	0.000862	0.000218	0.0003	0.000367	0	0.000105	0.000126	0.00187	2.87E-07	0.0004	0.000373	0.00027	0.00056	0.001981	0.000686	0.000658
ц	0.00333	1.23E-05	9.11E-05	9.74E-06	1.12E-05	0.000987	0	6.3E-06	6.59E-06	0.000134	0	1.75E-05	6.88E-06	3.41E-05	3.15E-05	0.000203	3.78E-05	1.03E-05
ч	0.001379	9.17E-06	3.21E-05	1.46E-05	1.98E-05	0.002792	8.6E-07	2.01E-06	4.3E-06	0.001817	0	6.22E-05	2.81E-05	4.56E-05	0.000487	7.62E-05	4.61E-05	0.000157
ш	0.001043	1.06E-05	2.03E-05	4.3E-06	8.31E-06	0.002158	6.3E-06	8.6E-07	1.72E-06	0.002	0	4.38E-05	0.000648	7.16E-06	0.000286	0.000123	1.78E-05	4.3E-06
щ	0.000521	5.73E-07	2.35E-05	5.73E-07	0	0.001519	8.6E-07	0	0	0.001208	0	2.87E-07	2.87E-07	0	5.87E-05	1.15E-06	5.73E-07	6.59E-06
ъ	0	0	0	0	0	4.07E-05	0	0	0	1.15E-06	3.72E-06	0	0	0	0	0	0	0
ы	6.22E-05	0.000447	0.001433	0.000201	0.000349	0.001476	0	9.83E-05	0.000175	0.001091	0.0001395	0.000408	0.001617	0.001136	0.002717	0.000459	0.000754	0.000311
ь	0.000137	0.000583	0.001121	0.000306	0.000602	0.000719	0	0.000121	0.000289	0.001527	0	0.000686	0.000226	0.000946	0.001538	0.000601	0.000894	0.000142
э	0	2.87E-07	8.6E-07	1.72E-06	5.73E-07	0	0	0	0	5.73E-07	5.73E-07	5.73E-07	8.31E-06	1.72E-06	1.15E-06	2.87E-07	1.15E-06	4.3E-06
ю	6.82E-05	0.000404	0.000444	0.000184	0.000583	0.000127	0	8.31E-05	0.000136	0.000992	1.15E-06	0.000225	5.13E-05	0.000138	0.000347	0.000234	0.000362	0.000105
я	0.000173	0.00047	0.001616	0.000534	0.000589	0.000621	0	0.000227	0.000645	0.001929	0.000104	0.000894	0.000853	0.001015	0.002095	0.000736	0.001146	0.000255
	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

р	с	т	у	ф	х	ц	ч	ш	щ	ъ	ы	ь	э	ю	я	
0.005337	0.00519	0.004473	0.000536	0.000419	0.001513	0.000229	0.001036	0.001346	0.000268	0	0	0	9.28E-05	0.001104	0.000965	0
0.001666	6.5E-05	3.47E-05	0.001967	0	1.09E-05	5.44E-06	2.01E-06	1.06E-05	8.34E-05	8.62E-05	0.003559	5.01E-05	2.46E-05	1.15E-06	0.00089	0
0.001561	0.004094	0.001116	0.001202	4.81E-05	0.000181	0.000182	0.000278	0.000692	2.58E-06	0	0.002773	0.000436	8.71E-05	8.31E-06	0.001066	0
0.000864	9.89E-05	9.25E-05	0.000538	1.72E-06	2.01E-06	7.16E-06	2.92E-05	1.6E-05	0	0	0	0	1.15E-06	5.73E-07	6.02E-06	0
0.00103	0.000542	0.000235	0.002738	7.45E-06	2.64E-05	0.000737	3.38E-05	9.11E-05	1.89E-05	5.44E-06	0.000917	0.001391	6.59E-06	3.44E-06	0.000558	0
0.006369	0.007809	0.00699	0.000629	0.000425	0.001145	0.00055	0.001725	0.001034	0.000429	0	0	0	0.000119	0.00066	0.000751	0
3.15E-06	7.16E-06	9.46E-06	8.6E-07	0	2.01E-06	0	1.63E-05	2.87E-07	2.87E-07	0	0	0	4.58E-06	5.73E-07	0	0
6.53E-05	1.38E-05	2.01E-06	0.000237	0	2.87E-07	0	2.38E-05	0	0	0	0	4.44E-05	5.73E-07	0	5.73E-07	0
0.001216	0.000279	9.25E-05	0.000334	8.88E-06	1.75E-05	1.4E-05	4.1E-05	5.16E-06	0	1E-05	0.000474	0.000224	1.66E-05	1.17E-05	0.00036	0
0.001849	0.009601	0.006665	0.001503	0.000321	0.004274	0.001106	0.000987	0.000778	0.000532	0	0	0	0.000123	0.000478	0.002487	0
0.000139	0.001399	0.000768	0.000127	1.38E-05	6.1E-05	0.000123	0.000159	0.00015	2.87E-06	0	0	0	2.03E-05	2.58E-06	9E-05	0
0.001846	0.000556	0.000925	0.000996	2.06E-05	4.41E-05	6.39E-05	0.000149	2.06E-05	1.15E-06	0	0	0	2.18E-05	1.4E-05	5.67E-05	0
0.000106	0.00144	0.000376	0.001514	3.61E-05	7.71E-05	0.000136	0.000248	2.78E-05	2.87E-06	0	0.000646	0.002113	5.62E-05	0.001292	0.002302	0
0.000243	0.001081	0.000442	0.00332	0.000488	6.39E-05	0.00013	0.000354	1.32E-05	2.32E-05	2.29E-06	0.000884	0.000212	3.55E-05	4.3E-06	0.000783	0
0.00011	0.000738	0.000238	0.001493	2.72E-05	4.87E-05	0.000311	0.000216	2.12E-05	0.000109	2.87E-07	0.003324	0.001011	1.32E-05	0.000155	0.00187	0
0.007602	0.010857	0.01007	0.000521	0.000271	0.000576	0.000299	0.00171	0.000957	0.000196	0	0	0	0.000166	0.001493	0.000964	0
0.007665	0.00085	0.000201	0.000914	9.74E-06	5.73E-07	1.46E-05	9.46E-06	3.12E-05	3.15E-06	0	0.000156	7.56E-05	2.87E-07	0	0.000368	0
0.000123	0.000717	0.000924	0.002626	3.27E-05	0.000121	0.000169	6.3E-05	0.000268	2.38E-05	8.6E-07	0.001904	0.000636	7.45E-06	0.000269	0.001019	0
0.00045	0.000986	0.011351	0.00139	8.62E-05	0.000725	0.00013	0.000232	6.19E-05	1.15E-06	3.18E-05	0.001831	0.002039	1.6E-05	0.000109	0.000404	0
0.002701	0.002438	0.000618	0.001186	2.64E-05	8.42E-05	0.000499	0.000662	3.32E-05	2.44E-05	1.15E-06	0.002384	0.005652	5.16E-05	1.95E-05	0.000738	0
0.000326	0.002924	0.002125	9.94E-05	6.02E-05	0.000615	0.000118	0.001203	0.000819	0.000358	0	0	0	2.15E-05	0.000818	0.000129	0
0.000105	6.91E-05	2.46E-05	6.85E-05	4.47E-05	5.73E-06	5.16E-06	4.01E-06	2.58E-06	0	0	2.61E-05	5.79E-05	2.01E-06	2.87E-07	1.46E-05	0
0.000658	0.000667	0.000275	0.000322	8.6E-06	4.04E-05	0.000107	0.000173	2.18E-05	3.44E-06	0	0	2.87E-07	1.26E-05	3.72E-06	6.42E-05	0
1.03E-05	2.44E-05	2.49E-05	0.000203	2.58E-06	2.01E-06	5.16E-06	5.73E-06	1.43E-06	0	0	0.000288	0	8.6E-07	0	1.43E-06	0
0.000157	6.07E-05	0.00293	0.000287	2.87E-07	2.58E-06	5.44E-06	2.72E-05	8.83E-05	0	0	0	0.000265	5.73E-07	5.73E-07	7.16E-06	0
4.3E-06	1.35E-05	9.74E-06	0.00025	1.72E-06	2.29E-06	2.49E-05	4.3E-06	3.72E-06	0	0	0	0.000815	0	0	0	0
6.59E-06	2.29E-06	2.29E-06	0.000179	0	2.87E-07	0	0	0	0	0	0	3.21E-05	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	9.25E-05	0
0.000311	0.001228	0.00104	0.000177	9.17E-06	0.001424	4.07E-05	0.000238	0.000758	4.24E-05	0	0	0	1.81E-05	4.3E-06	5.96E-05	0
0.000142	0.001898	0.000755	0.000201	1.89E-05	8.31E-05	0.000189	0.000253	0.000212	2.01E-05	0	0	0	4.64E-05	0.000376	0.000566	0
4.3E-06	5.73E-07	0.000994	0	1.15E-06	8.6E-07	1.15E-06	0	8.6E-07	0	0	0	0	5.73E-07	0	0	0
0.000105	0.000472	0.000952	5.64E-05	7.16E-06	3.24E-05	5.04E-05	0.000215	7.74E-06	0.000633	0	0	0	1.66E-05	3.3E-05	6.85E-05	0
0.000255	0.001377	0.002636	0.00026	2.49E-05	0.000281	0.000236	0.00055	3.64E-05	0.000778	0	0	0	3.61E-05	0.000166	0.000137	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Неперетинаючі біграми без пробілу:

	а	б	в	г	д	е	ё	ж	з	и	й	к	л	м	н	о	п	р
а	0.000723	0.001546	0.006532	0.001813	0.00312	0.001999	4.01E-06	0.001316	0.00388	0.00383	0.000951	0.004858	0.007732	0.005015	0.005272	0.001147	0.002181	0.00526
б	0.000539	7.45E-05	6.02E-05	8.6E-06	1.38E-05	0.002289	0	5.16E-06	5.16E-06	0.000961	0	2.41E-05	0.0015	4.01E-05	0.00034	0.00424	9.17E-06	0.001709
в	0.006263	0.000197	0.000649	0.000391	0.000797	0.007026	5.73E-06	0.000155	0.000787	0.004912	1.15E-06	0.00047	0.00107	0.000479	0.001242	0.01175	0.000922	0.001558
г	0.001321	4.18E-05	0.000113	5.44E-05	0.001378	0.000393	0	9.74E-06	2.92E-05	0.001066	0	4.58E-05	0.00105	4.53E-05	0.000664	0.014163	7.39E-05	0.000859
д	0.006539	8.31E-05	0.000997	0.000171	0.000119	0.007557	5.16E-06	3.32E-05	6.25E-05	0.00347	0	0.000167	0.001095	0.000175	0.002538	0.0043	0.00018	0.001083
е	0.000218	0.003742	0.004637	0.006176	0.004267	0.002003	4.01E-06	0.001291	0.001951	0.002922	0.002822	0.002058	0.00718	0.007237	0.009457	0.001365	0.002697	0.006305
ё	0	2.87E-06	3.44E-06	1.72E-06	3.44E-06	2.29E-06	0	1.15E-06	2.29E-06	6.3E-06	0	1.15E-06	9.74E-06	8.02E-06	9.17E-06	2.29E-06	6.3E-06	3.44E-06
ж	0.000776	4.01E-05	7.45E-06	2.46E-05	0.001266	0.003605	1.15E-06	0.00012	5.73E-07	0.00222	0	2.92E-05	2.87E-06	5.16E-06	0.000469	2.52E-05	4.01E-06	5.96E-05
з	0.004841	0.000323	0.001223	0.000266	0.000918	0.001353	0	4.24E-05	0.000284	0.000487	0	0.000165	0.000623	0.000285	0.001516	0.000696	0.000161	0.001176
и	0.000847	0.002791	0.005759	0.001189	0.003438	0.005261	0	0.000705	0.004272	0.004704	0.001371	0.003715	0.0071	0.00586	0.006048	0.003418	0.004197	0.001877
й	0.000121	0.000238	0.000597	0.000232	0.00096	0.000181	0	0.000124	0.000218	0.001256	0	0.000324	7.51E-05	0.00033	0.00073	0.000321	0.000626	0.000134
к	0.006085	0.000235	0.000516	0.000335	0.000189	0.000447	0	0.000198	7.85E-05	0.003027	0	0.000343	0.000805	0.000181	0.000816	0.008173	0.000422	0.001862
л	0.00626	0.0003	0.000755	0.000476	0.000365	0.005956	5.73E-07	0.000411	0.000117	0.010504	0	0.000787	0.000165	0.000531	0.00097	0.006049	0.00047	0.000111
м	0.002234	0.000382	0.000884	0.000383	0.00041	0.004805	0	9.46E-05	0.000168	0.005737	0	0.000674	0.00135	0.000441	0.002412	0.004401	0.000824	0.000272
н	0.011917	0.000284	0.000397	0.000324	0.000219	0.010103	3.44E-06	7.62E-05	9.05E-05	0.011568	0	0.000269	3.72E-05	0.000187	0.002652	0.007508	0.000452	0.000106
о	0.00506	0.004597	0.011146	0.006621	0.009395	0.003821	4.58E-06	0.002071	0.00208	0.005234	0.002655	0.002867	0.004703	0.005662	0.007534	0.001427	0.002597	0.007544
п	0.001623	1.72E-06	5.16E-06	4.01E-06	1.72E-06	0.001696	1.15E-06	5.73E-07	2.29E-06	0.000987	0	5.85E-05	0.00092	0	8.94E-05	0.011439	3.67E-05	0.007701
р	0.008423	0.000166	0.000477	0.000301	0.000719	0.007364	1.15E-06	0.000245	0.000182	0.006915	0	0.000224	0.000113	0.000197	0.000696	0.007606	0.000484	0.000136
с	0.002397	0.000183	0.003523	0.000175	0.000813	0.005665	6.19E-05	9.23E-05	8.54E-05	0.0031	0	0.00435	0.003966	0.000809	0.001404	0.003127	0.004342	0.000451
т	0.005027	0.000347	0.000562	0.000612	0.000715	0.006155	0	0.000135	0.000223	0.004822	0	0.000905	0.000332	0.000489	0.001372	0.012389	0.000979	0.002712
у	0.000124	0.000859	0.001159	0.001133	0.003968	0.000399	0	0.001148	0.000381	0.001457	0.000135	0.00126	0.000466	0.001239	0.000846	0.000317	0.001327	0.000325
ф	0.000619	3.15E-05	3.95E-05	2.81E-05	2.46E-05	0.000253	0	3.44E-06	1.09E-05	0.000457	0	4.07E-05	0.000127	0.000111	4.93E-05	0.000154	4.99E-05	0.000109
х	0.000716	0.000257	0.000838	0.000221	0.000296	0.000367	0	0.000111	0.000125	0.001811	5.73E-07	0.000403	0.000371	0.000273	0.000554	0.001954	0.00067	0.000666
ц	0.003368	1.09E-05	9E-05	1.2E-05	9.17E-06	0.000985	0	8.02E-06	1.03E-05	0.000122	0	1.89E-05	8.02E-06	4.01E-05	2.64E-05	0.000205	3.95E-05	8.6E-06
ч	0.001387	1.03E-05	3.21E-05	1.43E-05	1.49E-05	0.002823	1.72E-06	1.72E-06	4.58E-06	0.001838	0	5.9E-05	9.29E-05	4.93E-05	0.000473	7.28E-05	4.53E-05	0.00016
ш	0.001061	1.03E-05	2.64E-05	5.16E-06	9.17E-06	0.002171	4.58E-06	1.72E-06	2.29E-06	0.002031	0	4.36E-05	0.000632	8.6E-06	0.000287	0.000122	1.6E-05	3.44E-06
щ	0.000501	1.15E-06	2.01E-05	5.73E-07	0	0.001544	1.72E-06	0	0	0.001194	0	0	5.73E-07	0	5.33E-05	5.73E-07	1.15E-06	6.88E-06
ъ	0	0	0	0	0	4.36E-05	0	0	0	0	5.73E-07	5.16E-06	0	0	0	0	0	0
ы	6.76E-05	0.000442	0.00145	0.000195	0.000338	0.001476	0	9.11E-05	0.000189	0.001106	0.001402	0.000413	0.00157	0.001133	0.002751	0.000453	0.000759	0.000309
ь	0.000131	0.000569	0.001144	0.000029	0.000617	0.000723	0	0.000113	0.000294	0.001534	0	0.000696	0.000228	0.000931	0.001516	0.000596	0.000893	0.000144
э	0	0	1.15E-06	1.15E-06	0	0	0	0	0	0	5.73E-07	0	1.15E-06	8.6E-06	2.87E-06	0	5.73E-07	1.15E-06
ю	7.51E-05	0.000426	0.00044	0.000183	0.000582	0.000131	0	8.14E-05	0.000131	0.000951	1.72E-06	0.000223	5.16E-05	0.000136	0.00036	0.000237	0.000378	9.97E-04
я	0.00017	0.000477	0.001634	0.000537	0.000599	0.00063	0	0.000213	0.000648	0.001897	9.97E-05	0.000899	0.000859	0.001006	0.002114	0.000715	0.001154	0.00025
	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

р	с	т	у	ф	х	ц	ч	ш	щ	ъ	ы	ь	э	ю	я	
0.00526	0.005152	0.004474	0.000539	0.00042	0.001555	0.00023	0.001034	0.001331	0.00028	0	0	0	9E-05	0.001079	0.000989	0
0.001709	6.07E-05	3.21E-05	0.001951	0	9.74E-06	5.16E-06	2.87E-06	8.02E-06	7.68E-05	8.48E-05	0.003519	5.33E-05	3.04E-05	1.15E-06	0.000878	0
0.001558	0.004077	0.001136	0.001176	4.7E-05	0.000174	0.000018	0.00027	0.000684	2.29E-06	0	0.002772	0.000446	9.23E-05	9.17E-06	0.001082	0
0.000859	9.63E-05	0.0001	0.000527	2.29E-06	2.87E-06	7.45E-06	2.81E-05	1.26E-05	0	0	0	0	1.15E-06	0	5.73E-06	0
0.001083	0.000547	0.000246	0.002717	8.02E-06	2.98E-05	0.000739	3.5E-05	8.94E-05	1.83E-05	6.88E-06	0.000908	0.001417	8.02E-06	3.44E-06	0.00056	0
0.006305	0.007805	0.007029	0.000605	0.000413	0.001118	0.000543	0.001713	0.001041	0.000421	0	0	0	0.000123	0.000665	0.000732	0
3.44E-06	6.3E-06	1.03E-05	1.15E-06	0	2.87E-06	0	1.49E-05	0	5.73E-07	0	0	0	5.73E-06	1.15E-06	0	0
5.96E-05	1.43E-05	2.29E-06	0.000247	0	5.73E-07	0	2.12E-05	0	0	0	0	4.64E-05	0	0	0	0
0.001176	0.000291	8.88E-05	0.000343	1.09E-05	2.06E-05	1.66E-05	4.36E-05	5.73E-06	0	9.17E-06	0.000462	0.000214	1.83E-05	1.09E-05	0.000368	0
0.001877	0.009601	0.006665	0.001552	0.000313	0.004316	0.001109	0.000993	0.000752	0.000536	0	0	0	0.000121	0.000492	0.002489	0
0.000134	0.001417	0.000748	0.000144	1.38E-05	6.76E-05	0.00012	0.000148	0.000146	2.29E-06	0	0	0	2.23E-05	3.44E-06	8.65E-05	0
0.001862	0.000545	0.000933	0.00101	2.41E-05	3.78E-05	6.36E-05	0.000155	1.83E-05	0	0	0	0	1.78E-05	1.49E-05	5.56E-05	0
0.000111	0.001441	0.000385	0.001516	3.78E-05	7.56E-05	0.000134	0.000253	3.04E-05	4.01E-06	0	0.000658	0.00209	4.93E-05	0.001293	0.002265	0
0.000272	0.001076	0.00045	0.003314	0.000488	6.13E-05	0.000134	0.000343	1.6E-05	2.64E-05	1.72E-06	0.000885	0.000214	3.55E-05	5.73E-06	0.000755	0
0.000106	0.000719	0.000237	0.001513	2.81E-05	5.21E-05	0.000297	0.000225	1.83E-05	0.000107	0	0.003309	0.001006	1.38E-05	0.000156	0.001859	0
0.007544	0.010872	0.010133	0.000538	0.000265	0.000585	0.000285	0.001697	0.000981	0.000203	0	0	0	0.000168	0.001505	0.000949	0
0.007701	0.000851	0.000202	0.000885	1.32E-05	5.73E-07	1.78E-05	8.6E-06	3.15E-05	3.44E-06	0	0.00015	7.05E-05	0	0	0.000352	0
0.000136	0.000687	0.000943	0.002613	2.87E-05	0.000121	0.000173	7.39E-05	0.00027	2.64E-05	5.73E-07	0.00188	0.000656	9.17E-06	0.000261	0.001066	0
0.000451	0.001006	0.011292	0.001415	8.14E-05	0.000716	0.000127	0.000211	6.13E-05	5.73E-07	2.98E-05	0.001814	0.00205	1.72E-05	0.000116	0.004024	0
0.002712	0.002434	0.000638	0.001147	2.64E-05	9.11E-05	0.000491	0.000666	3.04E-05	2.46E-05	1.15E-06	0.002405	0.005668	5.56E-05	1.55E-05	0.000747	0
0.000325	0.002963	0.002166	0.000109	5.56E-05	0.000626	0.000116	0.001187	0.000798	0.000356	0	0	0	2.46E-05	0.000811	0.000116	0
0.000109	8.02E-05	2.41E-05	6.36E-05	4.7E-05	5.16E-06	5.73E-06	5.16E-06	2.87E-06	0	0	2.98E-05	5.96E-05	5.73E-07	0	1.2E-05	0
0.000666	0.000676	0.000268	0.000325	8.6E-06	4.47E-05	0.000109	0.000171	2.35E-05	1.15E-06	0	0	0	1.32E-05	2.87E-06	6.65E-05	0
8.6E-06	2.81E-05	2.75E-05	0.000205	2.29E-06	1.15E-06	5.73E-06	4.01E-06	1.72E-06	0	0	0.0003	0	5.73E-07	0	1.72E-06	0
0.00016	6.59E-05	0.002936	0.000288	0	2.87E-06	6.3E-06	2.69E-05	8.31E-05	0	0	0	0.000272	1.15E-06	0	5.73E-06	0
3.44E-06	9.17E-06	9.74E-06	0.000255	1.15E-06	1.72E-06	2.41E-05	5.16E-06	4.01E-06	0	0	0	0.0008	0	0	0	0
6.88E-06	2.87E-06	1.72E-06	0.000178	0	0	0	0	0	0	0	0	0	3.67E-05	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0.000309	0.001218	0.001053	0.000181	9.74E-06	0.00145	3.78E-05	0.000252	0.000772	4.13E-05	0	0	0	1.66E-05	4.58E-06	6.25E-05	0
0.000144	0.001893	0.000742	0.000197	2.06E-05	8.08E-05	0.000183	0.000249	0.000219	2.12E-05	0	0	0	4.18E-05	0.000377	0.000566	0
4.01E-06	5.73E-07	0.000985	0	5.73E-07	1.15E-06	5.73E-07	0	5.73E-07	0	0	0	0	1.15E-06	0	0	0
9.97E-05	0.000476	0.000959	5.44E-05	8.6E-06	3.04E-05	4.36E-05	0.000211	8.6E-06	0.000641	0	0	0	1.43E-05	3.15E-05	6.82E-05	0
0.00025	0.001381	0.002652	0.000264	2.52E-05	0.000299	0.000234	0.000538	3.38E-05	0.000778	0	0	0	3.9E-05	0.000168	0.000135	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Функція для обчислення ентропії:

```
def entropy(freq_dict, bigram=False):  
    H = -sum(p * math.log2(p) for p in freq_dict.values() if p > 0)  
    if bigram:  
        H /= 2  
    return H
```

Отримані значення ентропії

Ентропія

H1 букви з пробілом: 4.326846

H1 букви без пробілу: 4.438455

H2 перетинаючі біграми: 3.927804

H2 перетинаючі без пробілу: 4.134806

H2 неперетинаючі біграми: 3.928026

H2 неперетинаючі без пробілу: 4.134592

За допомогою програми CoolPinkProgram оцінили значення для $H(10)$, $H(20)$, $H(30)$:

Для 10 символів:

Лабораторная работа №1

Произвольная часть текста:
огласны_точно_так_же_не_имело_бы_смысла_говорить_что_футбольный_игрок_допус

Использованные буквы:
а, е, в,

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

Введенный символ: о

Символ по счету: 4

Номер эксперимента: 50

Неравенство для энтропии:
 $1,87546583878224 < H < 2,7777951914117$

Двоичная таблица угаданных символов:
01000000000000000000000000000000
01000000000000000000000000000000
10000000000000000000000000000000
10000000000000000000000000000000
10000000000000000000000000000000

Поле ввода символов:
о

Продолжить Другой

Вероятности:
q[1] = 0,46
q[2] = 0,14
q[3] = 0,1
q[4] = 0,04
q[5] = 0,04
q[6] = 0,04
q[7] = 0,02
q[8] = 0
q[9] = 0,02
q[10] = 0,04
q[11] = 0,02
q[12] = 0,02
q[13] = 0
q[14] = 0,02
q[15] = 0
q[16] = 0
q[17] = 0
q[18] = 0,02
q[19] = 0
q[20] = 0
q[21] = 0
q[22] = 0,02
q[23] = 0
q[24] = 0
q[25] = 0
q[26] = 0
q[27] = 0
q[28] = 0
q[29] = 0
q[30] = 0
q[31] = 0
q[32] = 0

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

Для 20 символов:

Лабораторная работа №1

Произвольная часть текста:

ой_нибудь_страны_могут_утверждать_что_договоры_не_имеют_никакого_значения_н

Использованные буквы:

Порядок n-граммы:

5 символов

10 символов

15 символов

20 символов

25 символов

30 символов

35 символов

40 символов

45 символов

50 символов

Введенный символ:

г

Символ по счету:

1

Номер эксперимента:

52

Неравенство для энтропии:

$2.69560895485705 < H < 3.42277506626134$

Двоичная таблица угаданных символов:

00000000000100000000000000000000
01000000000000000000000000000000
00000000000001000000000000000000
00000000000000000000000000000000
00000000000010000000000000000000
00000000000000000000000000000000

Поле ввода символов:

г

Продолжить

Другой

Вероятности:

q[1] = 0.3846153
q[2] = 0.1538461
q[3] = 0.0192307
q[4] = 0
q[5] = 0.0192307
q[6] = 0.0384615
q[7] = 0.0192307
q[8] = 0.0192307
q[9] = 0
q[10] = 0.0192307
q[11] = 0
q[12] = 0.0192307
q[13] = 0.0384615
q[14] = 0.0192307
q[15] = 0.0192307
q[16] = 0.0192307
q[17] = 0.0192307
q[18] = 0.0192307
q[19] = 0
q[20] = 0
q[21] = 0.0192307
q[22] = 0.0384615
q[23] = 0
q[24] = 0.0192307
q[25] = 0.0192307
q[26] = 0
q[27] = 0
q[28] = 0.0384615
q[29] = 0.0192307
q[30] = 0
q[31] = 0
q[32] = 0.0192307

Строка состояния:

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

Для 30 символов:

Лабораторная работа №1

Произвольная часть текста:

_или_братом_то_если_бы_я_знал_как_они_способны_раздражать_человека_я_бы_не_

Использованные буквы:

Порядок n-граммы:

5 символов

10 символов

15 символов

20 символов

25 символов

30 символов

35 символов

40 символов

45 символов

50 символов

Введенный символ:

_ (пробел)

Символ по счету:

1

Номер эксперимента:

55

Неравенство для энтропии:

$2.14939306272318 < H < 2.69327338391587$

Двоичная таблица угаданных символов:

10000000000000000000000000000000
10000000000000000000000000000000
10000000000000000000000000000000
10000000000000000000000000000000
00010000000000000000000000000000
00000000000000000000000000000000

Поле ввода символов:

Продолжить

Другой

Вероятности:

q[1] = 0.5636363
q[2] = 0.0181818
q[3] = 0.0181818
q[4] = 0.0363636
q[5] = 0.0181818
q[6] = 0.0181818
q[7] = 0.0181818
q[8] = 0.0545454
q[9] = 0.0181818
q[10] = 0
q[11] = 0
q[12] = 0
q[13] = 0
q[14] = 0.0181818
q[15] = 0
q[16] = 0
q[17] = 0.0181818
q[18] = 0
q[19] = 0.0181818
q[20] = 0.0181818
q[21] = 0
q[22] = 0
q[23] = 0
q[24] = 0
q[25] = 0.0181818
q[26] = 0.0181818
q[27] = 0.0545454
q[28] = 0.0545454
q[29] = 0
q[30] = 0.0181818
q[31] = 0
q[32] = 0

Строка состояния:

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

Отримали такі результати:

	Ентропія
H(10)	1,875<H<2,777
H(20)	2,695<H<3,422
H(30)	2,149<H<2,693

Також обрахували надлишковість в російській мові для кожного випадку. Для цього використали наступну формулу:

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

Функція для обчислення надлишковості:

```
def redundancy(H, max_H):  
    return 1 - (H / max_H)
```

Отримані дані з оцінками надлишковості:

```
Надлишковість  
R букви з пробілом: 0.149508  
R букви без пробілу: 0.120121  
R перетинаючі біграми: 0.227944  
R перетинаючі без пробілу: 0.180317  
R неперетинаючі біграми: 0.227901  
R неперетинаючі без пробілу: 0.180359
```

І також оцінки надлишковості для CoolPinkProgram:

	Надлишковість
H(10)	0,625<R<0,444
H(20)	0,461<R<0,315
H(30)	0,570<R<0,461

Висновки:

У результаті практикуму ми на практиці розібралися з поняттями ентропії та надлишковості, обчислили частоти символів і n-грам та оцінили їхній вплив на інформаційні характеристики тексту. За допомогою програми CoolPinkProgram

наближено визначили значення ентропії для різних довжин контексту. Це дало можливість побачити, як зі збільшенням кількості врахованих символів змінюється ентропія і, відповідно, надлишковість. Отримані результати допомогли краще зрозуміти властивості мови та закріпити навички її аналізу.