# МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ УКРАЇНИ «КИЇВСЬКИЙ ПОЛІТЕХНІЧНИЙ ІНСТИТУТ імені Ігоря Сікорського» «ФІЗИКО-ТЕХНІЧНИЙ ІНСТИТУТ»

#### КРИПТОГРАФІЯ

Комп'ютерний практикум №1

Виконали: студенти 3-го курсу

групи ФБ-22

Мартинюк А.В. та Шеїна Е.О.

Бригада №1

# **Експериментальна оцінка ентропії на символ джерела відкритого** тексту

#### Мета роботи

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

#### Порядок виконання роботи

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

#### Хід роботи

Для початку визначаємо алфавіт, що буде та не буде містити пробіл. chars\_with\_space = 'абвгдеёжзийклмнопрстуфхцчшщъыьэюя ' chars\_without\_space = 'абвгдеёжзийклмнопрстуфхцчшщъыьэюя' Вкажемо шлях до текстового файлу. def read\_text(file\_path):

```
with open(file path, 'r', encoding='utf-8') as file:
          return file.read().replace('n', ").replace('\n', '').lower()
Фільтруємо текст, залишаючи лише символи з алфавіту.
     def filter text(input text, charset, with space=False):
       if with space:
          return ".join(symbol for symbol in input text if symbol in
     charset and (symbol.isalpha() or symbol.isspace()))
       else:
          return ".join(symbol for symbol in input text if symbol in
     charset and symbol.isalpha())
Рахуємо кількість та частоту символів.
     def calculate char count and freq(txt: str):
       count dict = \{\}
       for char in txt:
          count dict[char] = count dict.get(char, 0) + 1
       total length = len(txt)
       freq dict = {char: round(count / total length, 5) for char, count in
     count dict.items()}
       return count dict, freq dict
Частоту рахували як відношення кількості певних символів до
відношення загальної кількості символів.
Створюємо таблицю символів.
     def create char freq df(count dict, freq dict):
       return pd.DataFrame({
          'Char': freq dict.keys(),
          'Count': [count dict[c] for c in freq dict.keys()],
          'Freq': freq dict.values()
```

```
}).sort values(by='Freq', ascending=False)
Рахуємо біграми.
     def calculate bigram count and freq(txt: str, overlap: bool):
       bigram count = {}
       step = 1 if overlap else 2
       for i in range(0, len(txt) - 1, step):
          bigram = txt[i] + txt[i + 1]
         bigram count[bigram] = bigram count.get(bigram, 0) + 1
       total bigrams = len(txt) - 1 if overlap else len(txt) // 2
       bigram freq = {bg: round(count / total bigrams, 5) for bg, count in
     bigram count.items()}
       return bigram count, bigram freq
Створюємо таблиці біграм.
     def create bigram df(bigram freq, charset):
       sorted charset = sorted(charset)
       bigram df = pd.DataFrame(index=sorted charset,
     columns=sorted charset)
       for bg in bigram freq:
         x = sorted charset.index(bg[0])
          y = sorted charset.index(bg[1])
          bigram df.iloc[x, y] = bigram freq[bg]
       return bigram df.rename(index={" ": "space"}, columns={" ":
     "space"}) if " " in bigram_df.index else bigram df
Рахуємо ентропію за формулою:
```

$$H(Z) = -\sum_{i=1}^{n} p_{i} \log p_{i}.$$

def compute\_entropy(freq\_dict):

return -sum(freq \* math.log2(freq) for freq in freq\_dict.values() if freq > 0)

Рахуємо надлишковість за формулою:

$$R = 1 - \frac{H_{\infty}}{H_{0}}.$$

def compute redundancy(h, charset):

h0 = math.log2(len(charset))

return 1 - (h / h0)

#### Отримуємо результат:

```
Ентропія Н1 для тексту без пробілів: 4.460131915026182
Ентропія Н1 для тексту з пробілами: 4.367174638930017
Ентропія Н2 для пересічних біграм (без пробілів): 8.284384402566635
Ентропія Н2 для пересічних біграм (з пробілами): 7.951525666621003
Ентропія Н2 для непересічних біграм (без пробілів): 8.275928012962769
Ентропія Н2 для непересічних біграм (з пробілами): 7.949713904306328
Надлишковість Н1 (без пробілів): 0.11582405944255969
Надлишковість Н1 (з пробілами): 0.1415810247261281
Надлишковість Н2 пересічні (без пробілів): -0.6422952304171357
Надлишковість Н2 пересічні (з пробілами): -0.5629648637722073
Надлишковість Н2 непересічні (без пробілів): -0.6406188368991483
Надлишковість Н2 непересічні (з пробілами): -0.5626087408144167
```

Монограми без пробілів:

Char	Count	Freq
0	19942	0,10984
e	15313	0,08435
a	14572	0,08026
и	12565	0,06921
н	12562	0,06919
т	12406	0,06833
С	9773	0,05383
p	8300	0,04572
л	8169	0,045
В	7379	0,04064
M	5960	0,03283
к	5642	0,03108
д	5318	0,02929
п	5083	0,028
у	4986	0,02746
я	3827	0,02108
ь	3596	0,01981
ы	3505	0,01931
Γ	3013	0,0166
3	3003	0,01654
6	2889	0,01591
ч	2833	0,0156
й	2148	0,01183
X	1902	0,01048
ж	1869	0,01029
ш	1671	0,0092
ю	1018	0,00561
щ	775	0,00427
ц	708	0,0039
Э	562	0,0031
ф	208	0,00115
ъ	54	0,0003

Монограми з пробілом:

Char	Count	Freq
	36247	0,16642
0	19942	0,09156
e	15313	0,07031
a	14572	0,06691
И	12565	0,05769
Н	12562	0,05768
Т	12406	0,05696
C	9773	0,04487
p	8300	0,03811
л	8169	0,03751
В	7379	0,03388
M	5960	0,02736
К	5642	0,0259
Д	5318	0,02442
п	5083	0,02334
у	4986	0,02289
Я	3827	0,01757
ь	3596	0,01651
ы	3505	0,01609
Γ	3013	0,01383
3	3003	0,01379
б	2889	0,01326
ч	2833	0,01301
й	2148	0,00986
x	1902	0,00873
ж	1869	0,00858
Ш	1671	0,00767
ю	1018	0,00467
щ	775	0,00356
ц	708	0,00325
3	562	0,00258
ф	208	0,00096
ъ	54	0,00025

# Пересічні біграми без пробілів:

6			0.00546																														
	0,00077			0,00128	0,00263	0,00188	0,00151	0,00404	0,00123	0,001	0,00588	0,00922	0,00465	0,00809	0,00107	0,00283	0,00536	0,00608	0,00619	0,0005	0,00013		0,00051	0,00158							0,00128		
8 1		0,00001	0,00007	0,00001	0,00003	0,00213	0,00001	0,00001	0,00093		0,00005	0,0007	0,00004	0,00033	0,0029	0,00002	0,00138	0,00019	0,00002	0,00126		0,00006			0,00002	0,00034	0,00023	0,00376	0,00004	0,00003	0,00001	0,00058	
	0,00596	0,00018	0,00028	0,00029	0,00095	0,00542	0,00005	0,00047	0,00318		0,00064	0,00095	0,00035	0,00175	0,00735	0,00103	0,00097	0,00467	0,00102	0,00078	0,00002	0,00005	0,00007	0,00023	0,00102	0,00001		0,00241	0,00012	0,00015	0,00002	0,00025	
r (	0,00155	0,00003	0,00009	0,00001	0,00159	0,00025	0,00001	0,00003	0,0009		0,00013	0,00154	0,00002	0,00048	0,00804	0,00009	0,00114	0,00009	0,00004	0,00049		0,00001	0,00001	0,00006								0,00001	
A 1	0,00545	0,00009	0,00082	0,00008	0,00012	0,00466	0,00005	0,00015	0,00295		0,0003	0,00074	0,00022	0,00248	0,00427	0,00024	0,00191	0,00041	0,00026	0,00169	0,00001	0,00003	0,00015	0,00007	0,00009		0,00003	0,00072	0,0007	0,00002	0,00003	0,00058	
e	0,0004	0,00248	0,00356	0,00398	0,00419	0,00149	0,00118	0,00275	0,00113	0,00237	0,00252	0,00614	0,00794	0,00935	0,00177	0,00351	0,00786	0,00681	0,00724	0,00073	0,0001	0,00116	0,00037	0,00179	0,00158	0,00124				0,00029	0,0001	0,0003	
ж	0,00113	0,00002	0,00001	0,00003	0,00074	0,00452	0,00003		0,002		0,00008	0,00003	0,00001	0,00104	0,00005	0,00001	0,00003	0,00023	0,00001	0,00028		0,00001		0,00004					0,00003				
3	0,00571	0,00017	0,00114	0,00052	0,00096	0,00046	0,00006	0,00008	0,00056		0,00025	0,00029	0,00046	0,00232	0,00087	0,00026	0,00042	0,0002	0,00018	0,00052	0,00001	0,00001		0,00003	0,00001		0,00001	0,00064	0,0001	0,00007	0,00001	0,00022	
и (	0,00083	0,00159	0,0049	0,00115	0,00296	0,003	0,00074	0,00283	0,00303	0,00139	0,00333	0,00531	0,00413	0,00639	0,00197	0,00298	0,00215	0,00548	0,00547	0,00065	0,00007	0,00259	0,0005	0,00215	0,00079	0,00021				0,00024	0,00047	0,00189	
й	0,00035	0,00028	0,00066	0,00019	0,00072	0,00011	0,0002	0,00025	0,00075		0,00059	0,00032	0,00045	0,0013	0,00054	0,001	0,00041	0,00158	0,0009	0,00019	0,00009	0,00006	0,00007	0,00045	0,00019	0,00001				0,00009	0,00002	0,00006	
K I	0,00696	0,00017	0,00052	0,00007	0,00018	0,00063	0,00015	0,00023	0,00367		0,00027	0,00052	0,00016	0,00095	0,00971	0,00035	0,00209	0,00074	0,00139	0,00177	0,00001	0,00004	0,00006	0,00026	0,00003					0,0001		0,00004	
л	0,00514	0,00026	0,00057	0,00035	0,00023	0,00502	0,00055	0,00024	0,00894		0,00061	0,00033	0,00014	0,00116	0,00643	0,00053	0,00014	0,00267	0,00042	0,00156	0,00001	0,00008	0,00002	0,00075	0,00004			0,00103	0,00451	0,00006	0,00122	0,00198	
M I	0,00365	0,00042	0,001	0,00031	0,00042	0,00442	0,00014	0,00036	0,00433		0,00058	0,00031	0,00052	0,00257	0,00449	0,00121	0,00037	0,0012	0,00066	0,00317	0,00003	0,00009	0,00006	0,00051	0,00007	0,00001		0,00127	0,00006	0,00012		0,0005	
н (	0,01222	0,00016	0,00021	0,00006	0,00098	0,01183	0,00003	0,00035	0,0102		0,00049	0,00003	0,00014	0,00458	0,01216	0,00028	0,00011	0,00073	0,00135	0,00308	0,0001	0,00022	0,0014	0,0002	0,00001	0,00006		0,0048	0,00165	0,00005	0,0001	0,00159	
0	0,00055	0,00558	0,00965	0,00551	0,0062	0,00319	0,00249	0,00221	0,00237	0,00502	0,00309	0,00655	0,00662	0,00939	0,00213	0,00442	0,00618	0,00984	0,01071	0,00067	0,00021	0,00105	0,00014	0,00292	0,001	0,00042				0,00056	0,00021	0,00097	
n (	0,00213	0,00001	0,00001		0,00002	0,00281		0,00001	0,00066		0,0001	0,00107		0,00017	0,01135	0,00007	0,00718	0,00004	0,00007	0,00136				0,00003	0,00001			0,00046	0,00002			0,00044	
р (	0,00892	0,00041	0,00055	0,00037	0,00032	0,0073	0,00033	0,00012	0,0049		0,00037	0,00013	0,0006	0,00157	0,00858	0,00021	0,00009	0,00036	0,0031	0,00276		0,00031	0,00003	0,00013	0,0002	0,00002		0,0017	0,00058	0,00001	0,00019	0,00156	
c 1	0,00185	0,0002	0,00206	0,00015	0,00038	0,00436	0,00009	0,00009	0,00162		0,00443	0,00367	0,00106	0,00199	0,00322	0,0028	0,00032	0,00112	0,01316	0,00097	0,00004	0,00023	0,00002	0,0003	0,0001	0,00001	0,00003	0,00031	0,00351	0,00008	0,00021	0,00546	
т (	0,00882	0,00018	0,00287	0,00014	0,00041	0,00793	0,00011	0,00026	0,00457		0,00128	0,00032	0,00032	0,0024	0,01623	0,00091	0,00479	0,00207	0,00066	0,00265	0,0001	0,00008	0,00004	0,0009	0,00007	0,00004		0,00209	0,00702	0,00014	0,00005	0,00088	
y	0,00025	0,00071	0,00112	0,00113	0,00217	0,00025	0,00178	0,00036	0,00059	0,0001	0,00117	0,00182	0,00126	0,00219	0,00039	0,0012	0,00121	0,00215	0,00261	0,00015	0,00005	0,0009	0,00003	0,00134	0,00077	0,00039				0,0001	0,00111	0,00016	
φ (	0,00019					0,00015			0,0002		0,00001	0,00006		0,00001	0,00044		0,00002			0,00003	0,00001							0,00001	0,00001				
× I	0,00102	0,00014	0,00048	0,00012	0,00019	0,00013	0,00004	0,00018	0,00063		0,00029	0,0003	0,0002	0,00075	0,00327	0,00064	0,00032	0,00076	0,00034	0,00027	0,00004	0,00002	0,00002	0,00014	0,00004		0,00001			0,00007	0,00001	0,00005	
ц (	0,00052	0,00001	0,00012		0,00003	0,00059		0,00002	0,00204		0,00004		0,00001	0,00003	0,00021	0,00003	0,00001	0,00004	0,00001	0,00009		0,00001		0,00001				0,00009		0,00001		0,00001	
м (	0,00242	0,00001	0,00003	0,00001	0,00002	0,00461		0,00001	0,00203		0,00029	0,00001	0,00001	0,00097	0,00013	0,00002	0,00001	0,00003	0,00397	0,00062	0,00002	0,00001			0,00014				0,00025	0,00001			
w (	0,00131	0,00001	0,00005	0,00001		0,00227			0,00215		0,00036	0,00057	0,00008	0,00052	0,00023	0,00003	0,00002	0,00001	0,00009	0,00034			0,00001	0,00001					0,00116				
щ	0,00069					0,00207			0,00112				0,00001	0,0001		0,00001	0,00001			0,00019				0,00001					0,00007				
ъ						0,00014																										0,00016	
ы	0,00009	0,00035	0,00164	0,00023	0,00037	0,00182	0,00015	0,00025	0,00046	0,00185	0,00047	0,00246	0,00172	0,00061	0,00038	0,00058	0,00037	0,00124	0,001	0,00018		0,00166	0,00002	0,00034	0,00084					0,00014	0,00001	0,00007	
ь (	0,00035	0,0005	0,0011	0,00019	0,00052	0,00059	0,00009	0,00057	0,00104		0,00211	0,0002	0,00057	0,00322	0,00078	0,00125	0,00032	0,00257	0,00085	0,00023	0,00006	0,00009	0,00019	0,0006	0,00094	0,00002				0,0002	0,00041	0,00026	
9			0,00003	0,00001						0,00003	0,0001	0,00008	0,00001	0,00003		0,00001		0,00004	0,00269		0,00002	0,00006											
ю	0,00009	0,00027	0,00019	0,00013	0,00072	0,00003	0,00014	0,0001	0,00015		0,00028	0,00007	0,00011	0,00023	0,00018	0,00027	0,00013	0,0004	0,00094	0,00004	0,00001	0,00002	0,00014	0,00021	0,00002	0,00062				0,00006	0,00002	0,00005	
я (	0,00047	0,0004	0,00151	0,0003	0,00124	0,00031	0,00035	0,00059	0,00086	0,00007	0,00095	0,00124	0,00101	0,00223	0,00072	0,00121	0,00039	0,00207	0,003	0,00026	0,00004	0,00041	0,00004	0,00056	0,00007	0,00032				0,00014	0,00013	0,00017	
ë																																	

Пересічні біграми з пробілами:

	space	a	6		г	А	e	ж	2	и	й	К	л	M	н	0	n	р	c	T	у	ф	x	ц	4	ш	щ	ъ	ы		•	ю	a ě
space	0,00809	0,00446	0,00573	0,013	0,00312	0,00636	0,00311	0,00198	0,00483	0,01029		0,00798	0,00242	0,00523	0,01702	0,01035	0,01631	0,00404	0,01594	0,00967	0,00388	0,0005	0,00129	0,0003	0,00589	0,00087	0,00006				0,00246	0,00016	
а	0,01419	0			0,00083	0,00164	0,0013	0,00111			0,00083	0,0041	0,00749	0,00339	0,0053	0,00002	0,00081	0,00411	0,00354	0,00423	0,00008	0,00005		0,00042	0,0009	0,00087						0,00102	
6	0,00012	0,00064		0,00005	0	0,00003	0,00177	0,00001	0	0,00076		0,00004	0,00058	0,00003	0,00026	0,00239		0,00115	0,00016	0,00001	0,00105		0,00005			0,00001	0,00028	0,00019	0,00314	0,00003		0,00001	0,00049
8	0,00477	0,00488	0,00002	0,00006	0,00005	0,00055	0,00445	0	0,00028	0,00241		0,00016	0,00072	0,00012	0,00107	0,0058	0,00029	0,00071	0,00331	0,00039	0,0006		0,00004	0,00004	0,00006	0,00082	0,00001		0,00201	0,0001			0,0002
r	0,00067	0,00127	0	0,00001		0,00128	0,00019			0,00072		0,00007	0,00128		0,00035	0,00664		0,00093	0	0,00001	0,00037				0,00003								
A	0,00116	0,00451	0,00004	0,00059	0,00001	0,00005	0,00387	0,00003	0,00004	0,0024		0,00017	0,00058	0,00015	0,00189	0,00353	0,00006	0,00157	0,00028	0,00016	0,00137		0,00003	0,00012	0,00005	0,00007		0,00003	0,0006	0,00058		0,00002	0,00048
e	0,0171	0,00004	0,00112	0,00163	0,00299	0,00278	0,00104	0,00076	0,00156	0,0001	0,00198	0,00135	0,00492	0,00593	0,00616	0,00019	0,00102	0,00602	0,00383	0,00518	0,00011	0,00003	0,00078	0,0003	0,00093	0,00119	0,00102					0,00007	0,00017
360	0,00008	0,00094	0,00001	0	0,00002	0,00062	0,00376	0,00002		0,00166		0,00006	0,00002	0	0,00085	0,00004		0,00002	0,00017		0,00023				0,00004					0,00002			
3	0,00155	0,00473	0,00011	0,00079	0,00035	0,00074	0,00036	0,00005	0,00005	0,00042		0,0001	0,00023	0,00033	0,0018	0,00063	0,00001	0,00032		0,00001	0,00039				0	0		0	0,00054	0,00009		0	0,00017
и	0,01957	0,00016	0,00061	0,00237	0,00047	0,00158	0,0021	0,00043	0,00184	0,00127	0,00116	0,00187	0,00416	0,00296	0,00316	0,00027	0,00027	0,00129	0,00264	0,00339	0,00001	0,00002	0,00199	0,00039	0,00117	0,00055	0,00017				0	0,00038	0,00144
й	0,0081					0,00025						0,00008	0,00002	0,00005	0,00045			0	0,00035	0,00032		0,00001	0,00001	0,00001	0,00014	0,00006							0
K	0,00397	0,00572		0,00013	0		0,00044	0,00001	0,00008	0,00273		0,00004	0,00039		0,00037	0,00786		0,00159	0,00022	0,00091	0,00138			0,00004	0,00001	0,00001							
л	0,00585	0,00385	0,00004	0,00002	0,00017	0,00004	0,00394	0,00033	0,00007	0,00712		0,00023	0,00017	0	0,00039	0,00486	0		0,00171	0,00009	0,00117				0,00021	0,00002			0,00086	0,00376		0,00102	0,00157
M	0,00896	0,00281	0,00005	0,00001	0,00006	0	0,00352	0	0,00001	0,00298		0,00005	0,00017	0,00011	0,00118	0,0032	0,00011	0,00008	0,00011	0,00011	0,00231	0		0,00002	0,00003		0		0,00106	0,00005			0,00037
н	0,00258	0,01014			0,00001	0,0007	0,00982	0,00001	0,00022	0,0083		0,00028			0,00353	0,00994	0	0,00001	0,00033	0,00092	0,00249	0,00009	0,00017	0,00116	0,00012	0	0,00005		0,004	0,00137		0,00008	0,00132
0	0,02169	0,00005	0,00358	0,00605	0,00417	0,0043	0,00217	0,0018	0,00114	0,00077	0,00419	0,00179	0,00508	0,00477	0,0055	0,00034	0,0015	0,00472	0,00607	0,00748	0,00004	0,0001	0,00067	0,00008	0,00153	0,00071	0,00034				0,00009	0,00018	0,00066
n	0,00009	0,00178					0,00234			0,00054		0,00008	0,0009		0,00012	0,00944	0,00006	0,00599	0,00004	0,00006	0,00113				0,00002	0			0,00039	0,00001			0,00037
р	0,00096	0,0073	0,00033	0,0004	0,00028	0,00025	0,00608	0,00026	0,00005	0,00399		0,00023	0,0001	0,00048	0,00119	0,00709	0,0001	0,00007	0,00021	0,00256	0,00228		0,00025	0,00002	0,00009	0,00017	0,00001		0,00142	0,00048		0,00016	0,0013
c	0,00286	0,00146	0,00008	0,0016	0,00005	0,00018	0,00359	0,00002	0,00002	0,0013		0,00343	0,00304	0,00081	0,00126	0,00248	0,00205	0,00022	0,00071	0,01065	0,00075	0	0,00015	0,00001	0,00017	0,00007		0,00002	0,00026	0,00292		0,00016	0,00454
T	0,00641	0,00716	0,00003	0,00185	0,00002	0,0001	0,00644	0	0,00001	0,00333		0,00079	0,00016	0,00005	0,00125	0,01315	0,00011	0,00387	0,00118	0,00009	0,00206	0,00008	0,00003	0,00002	0,00041	0,00002	0,00004		0,00174	0,00585		0,00003	0,00068
У	0,00579	0,00003	0,00046	0,00047	0,0009	0,00161	0,00011	0,00141	0,00017	0,00001	0,00009	0,0006	0,00145	0,00079	0,00094		0,00047	0,0009	0,00125	0,00187		0,00003	0,0007	0,00003	0,00082	0,00063	0,00032				0,00003	0,00092	0,0001
ф	0	0,00016					0,00013			0,00017		0	0,00005		0,00001	0,00037		0,00002			0,00003	0							0,00001	0,00001			
×	0,00427			0,00011	0		0,00003			0,00017			0,00015	0,00005	0,00028			0,00015	0,00009	0,00004					0	0,00001		0					
ц	0,00025	0,00041		0,00007		0	0,00048			0,0017		0,00003				0,00016					0,00006								0,00008				
4	0,0002	0,00201		0,00001			0,00384			0,00169		0,00021	0,00001	0	0,00079	0,0001	0	0,00001		0,0033	0,00051					0,00011				0,00021			
w	0,00007			0,00003			0,00189			0,0018		0,0003	0,00048	0,00007	0,00043	0,00018	0,00002	0		0,00007				0						0,00096			
щ	0,00002	0,00058					0,00172			0,00093					0,00009			0			0,00016									0,00006			
ъ							0,00011																										0,00013
ы	0,00425		0,00013	0,00098	0,00009	0,00014	0,00143	0,00007	0,00005	0,00001	0,00154	0,00022	0,00199	0,00131	0,0001		0,00009	0,00019	0,00065	0,00065	0			0,00001	0,00016	0,00069							0,00002
	0,01042		0,00007	0,00001	0,00003	0,00002	0,0002		0,00014	0,00006		0,00124		0,00014	0,00135	0,00002			0,00126	0,00011		0,00002		0,00013	0,00005	0,00074	0,00001					0,00034	0,00015
				0,00002	0						0,00002	0,00009	0,00006	0	0,00002		0,00001		0,00004	0,00225		0,00001	0,00005										
10	0,00215		0,00018		0,00007	0,0005		0,00007	0,00002			0,00011	0	0,00004	0,00004			0,00002	0,00009	0,00062			0,00002	0,00011	0,00011	0,00001	0,00051					0,00001	
я	0,01023		0,00001	0,00028	0,00012	0,00067	0,00007	0,00019	0,00024	0	0,00006	0,00021	0,00087	0,00056	0,00053		0,00005	0,00011	0,00071	0,00182	0		0,00024	0,00003	0,00008	0,00001	0,00027					0,00011	0,0001

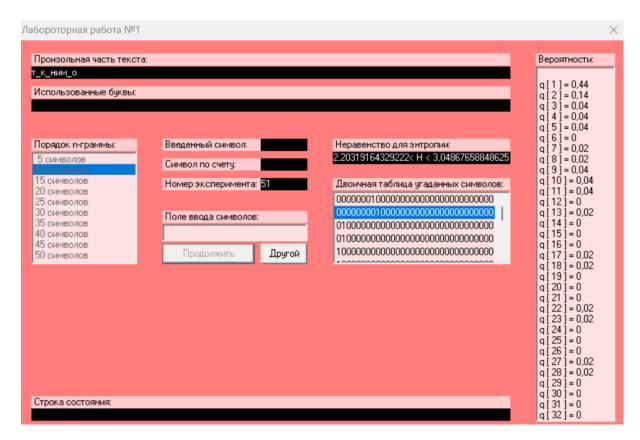
## Непересічні біграми без пробілів:

	a	6	в	r	д	e	ж	3	и	й	ĸ	л	M	н	0	n	р	c	т 1	y	Φ	x	ц	4	ш	щІ	ъ	ы	ь	9	10	я	ě
а	0,00046	0,00122	0,00549	0,00118	0,00264	0,00188	0,0015	0,00395	0,00123	0,00096	0,0056	0,00875	0,00464	0,00787	0,00109	0,00275	0,0057	0,00596	0,0059	0,00052	0,00017	0,00121	0,00062	0,00163	0,0012	0,00047				0,0003	0,00133	0,00288	
6	0,00067		0,00006		0,00003	0,00223	0,00001		0,00099		0,00003	0,00064	0,00007	0,00036	0,00308	0,00001	0,00132	0,00015	0,00002	0,00116		0,00002				0,00036	0,00025	0,00387	0,00002	0,00003	0,00002	0,00061	
8	0,0061	0,00018	0,0003	0,00029	0,00107	0,0051	0,00006	0,00055	0,00325		0,00065	0,00096	0,00034	0,00202	0,00729	0,00095	0,0009	0,0047	0,00099	0,00075	0,00003	0,00006	0,0001	0,00028	0,00119	0,00001		0,00252	0,00012	0,00014	0,00002	0,00024	
г	0,00159	0,00003	0,00012		0,00159	0,0003	0,00001	0,00004	0,00082		0,0001	0,00161	0,00001	0,00045	0,00796	0,00007	0,00107	0,0001	0,00004	0,00047		0,00001		0,00008								0,00001	
д	0,00541	0,00009	0,00079	0,00003	0,0001	0,00433	0,00006	0,00017	0,00305		0,00036	0,00078	0,00022	0,00236	0,00456	0,00026	0,00193	0,00046	0,00022	0,00173		0,00004	0,00015	0,00007	0,0001		0,00003	0,00087	0,00068	0,00001	0,00001	0,00059	
e	0,00037	0,00236	0,00347	0,00397	0,0042	0,00152	0,00107	0,0029	0,00109	0,00237	0,00258	0,00616	0,0078	0,00989	0,00166	0,00378	0,00789	0,00704	0,0073	0,00072	0,00004	0,00107	0,00037		0,00177	0,0013				0,00029	0,00014	0,00037	
ж	0,00129	0,00001		0,00003	0,00071		0,00004		0,00197		0,0001	0,00006	0,00001				0,00002	0,00025	0,00001	0,00033				0,00004					0,00002				
3	0,00582	0,0002	0,00119	0,00055	0,00078	0,00043	0,00009	0,00009	0,00058		0,00029	0,00028	0,00042	0,00226	0,00089	0,0002	0,00039	0,0002	0,00015	0,00058		0,00001		0,00002			0,00001	0,00057	0,00008	0,00003	0,00002	0,00023	
и	0,00066		0,00456		0,00291	0,00313		0,00283		0,00139	0,00332	0,00535	0,00381	0,00638		0,00288	0,00194		0,00544		0,00004	0,0026			0,00083					0,00029	0,00034		
й							0,00021	0,00018	0,00088		0,00064	0,00029	0,00043	0,0013				0,00163			0,00013	0,00004			0,00022	0,00001					0,00001		
K	0,00708	0,00012	0,00055	0,00004	0,00015	0,00072	0,00023	0,00029	0,00355		0,00028	0,00057	0,00015	0,0009	0,00997	0,00032	0,00205	0,00078	0,0015	0,00165		0,00002	0,00007	0,00032	0,00003					0,0001		0,00004	
л	0,00523	0,00028	0,00056	0,0004	0,00019		0,00052	0,00026	0,00917		0,00051	0,00039	0,00013	0,00098	0,00662	0,00056	0,00012	0,00274	0,0004	0,00164		0,00004	0,00001	0,00078	0,00004			0,00104	0,00429		0,00131		
M						0,00423			0,00442		0,00055		0,0005	0,00271			0,00036		0,00061		0,00003	0,00007	0,00011		0,00007			0,0013		0,00013		0,00047	
н	0,01205	0,00023	0,00023	0,00006	0,00096		0,00003	0,00031	0,01028		0,00042	0,00006	0,00018	0,00455	0,01208	0,00025	0,00013	0,00074	0,00135	0,00303	0,00011	0,00028	0,00143	0,00025	0,00001	0,00006		0,00458	0,00165	0,00007	0,00011	0,00175	
0	0,00052	0,00564	0,00945	0,00561		0,0031	0,0023			0,00508	0,00297	0,00646	0,00674	0,00912	0,0023	0,00447	0,00617	0,00971			0,00026	0,00097	0,00008		0,00106	0,00044				0,00054	0,0002	0,00089	
п	0,00218		0,00001					0,00001	0,00074		0,00008	0,00097				0,00008	0,00733	0,00004	0,0001	0,00135				0,00003				0,00044	0,00001			0,00036	
p	0,00945	0,00043					0,00033	0,00008	0,00492		0,00037	0,00004	0,00056	0,0015	0,00853	0,00019	0,0001	0,00037	0,00292	0,00284		0,00024	0,00003	0,00012	0,00021	0,00002		0,00176	0,00056		0,00021	0,00158	
C	0,00184	0,00015	0,00224	0,00014	0,00045			0,00009	0,00175		0,00421	0,00364	0,00098	0,00203	0,00326	0,0028	0,00021	0,001	0,01289	0,00095	0,00004	0,00023	0,00001	0,00035	0,00009		0,00004		0,00359	0,00004		0,00529	
		0,00014	0,00303	0,00014		0,00804		0,00031	0,00437		0,00137	0,00034	0,00034	0,00236		0,00096	0,00485	0,00208	0,00067	0,00281	0,00012	0,00013	0,00001	0,0009	0,00007	0,00004		0,0022	0,00722	0,00014		0,00089	
У		0,00074	0,00104	0,0012	0,00224		0,00188	0,00035		0,00017	0,00119		0,00132	0,00227	0,00039	0,00099	0,0012	0,0023	0,00242	0,00017	0,00007	0,00093	0,00001	0,00133	0,00067	0,0004				0,00011	0,00128	0,00017	
	0,00013					0,00018			0,00024			0,00003		0,00001	0,00047		0,00002			0,00003									0,00001				
x		0,00015	0,00052	0,00014		0,0001	0,00006	0,00018	0,00074		0,00028	0,00032	0,0002	0,00085			0,00037	0,00076	0,0003	0,00024			0,00001		0,00004		0,00001			0,00007		0,00004	
ц					0,00003			0,00002	0,002		0,00006		0,00001	0,00001		0,00002		0,00003		0,0001		0,00001		0,00001				0,00009				0,00001	
- 4	0,00243	0,00001		0,00001	0,00004				0,00199		0,00028		0,00001	0,00106		0,00003	0,00002	0,00002			0,00001	0,00001			0,00014				0,00026	0,00001			
ш	0,0013		0,00002			0,0021			0,00192		0,0004	0,00073			0,00018		0,00001	0,00002	0,00011	0,0003			0,00001						0,00106				
Щ	0,00076					0,00202			0,00118				0,00001	0,0001		0,00001				0,00014				0,00001					0,00011				
_ъ						0,0001																										0,00014	
ы	0,00011	0,0003	0,00158					0,0002		0,00184						0,00061		0,00122				0,00158								0,00012		0,00004	
b	0,0004	0,00043			0,00048	0,00054	0,00008	0,00059	0,0012		0,00214		0,00057		0,00066		0,00034	0,00265		0,00018			0,00021	0,00065	0,00101	0,00002				0,00024	0,00034	0,00031	
				0,00001										0,00002		0,00001		0,00006	0,00282			0,00003											
_ ю		0,00028	0,00019					0,00008	0,00013		0,00024	0,00002	0,0001	0,00026		0,00028			0,00104	0,00003	0,00002	0,00002	0,00009	0,00019						0,00009	0,00001		
8	0,00044	0,00043	0,00148	0,00031	0,00119	0,00026	0,00036	0,00056	0,00099	0,00008	0,00105	0,00116	0,0009	0,00208	0,00072	0,00116	0,00036	0,00202	0,00318	0,00023	0,00003	0,0004	0,00002	0,00059	0,00007	0,00032				0,00012	0,0001	0,00018	
ě																																	

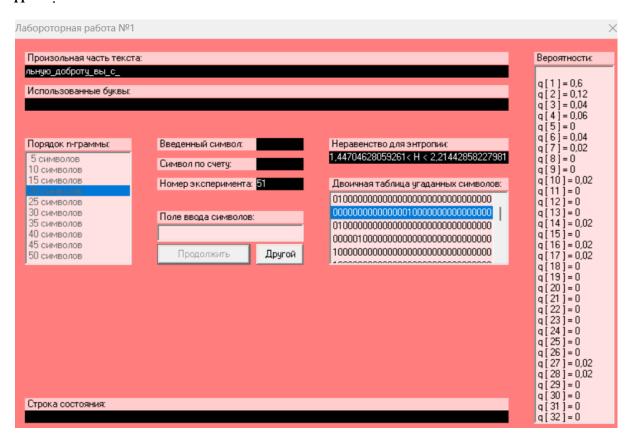
### Непересічні біграми з пробілами:

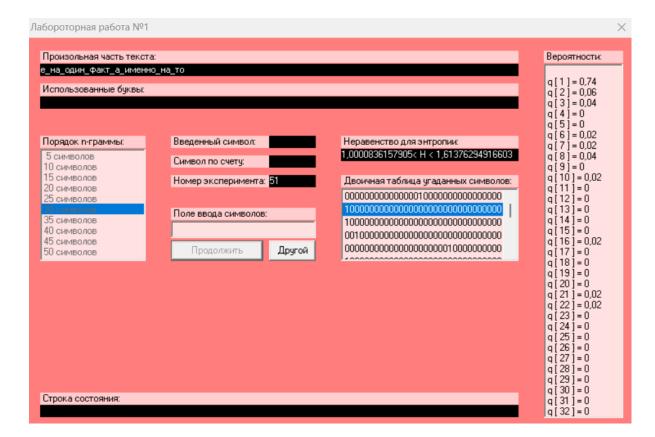
	space	3	6		r	A	e	ж	2	и	й	K	л	M	н	0	n	Р	c	T	y	ф	x	ц	ч	ш	щ	ъ	ы		2	ю	я	ě
space	0,00801	0,00402	0,00579	0,01326	0,00298	0,0063	0,00294	0,002	0,00488	0,01052		0,00784	0,00249	0,00521	0,01714	0,01008	0,01627	0,0041	0,01592	0,0098	0,00381	0,00047	0,00125	0,00026	0,00589	0,00089	0,00007				0,00253	0,00012	0,00086	
8	0,01436		0,00077	0,00327	0,00083	0,00175	0,00129	0,00103	0,00321	0,00012	0,00087	0,00405	0,00779	0,00332	0,00542	0,00002	0,00087	0,00416	0,00354	0,0041	0,00007	0,00005	0,00098	0,00042	0,00083	0,00093	0,00046					0,00108	0,0023	
6	0,0001	0,00064		0,00004		0,00002	0,00186		0,00001	0,00073		0,00004	0,0006	0,00002	0,00026	0,00229		0,00123	0,00017	0,00002	0,00111		0,00006			0,00002	0,00028	0,00018	0,0029	0,00003		0,00002	0,00036	
8	0,00455	0,00523	0,00004	0,00007	0,00002	0,00048	0,00424		0,00028	0,00244		0,00017	0,00066	0,00013	0,00096	0,00601	0,00034	0,00079	0,00311	0,00039	0,00062		0,00003	0,00003	0,00005	0,00088	0,00001		0,00192	0,00007			0,00025	
r	0,00057	0,00138		0,00001		0,00133	0,0002			0,00065		0,00008	0,0013		0,00032	0,00663		0,00099	0,00001	0,00002	0,00037				0,00001									
A	0,00116	0,00459	0,00005	0,00053	0,00001	0,00006	0,00365	0,00004	0,00002	0,00238		0,00016	0,00053	0,0001	0,00197	0,00351	0,00003	0,00182	0,00022	0,00014	0,00137		0,00005	0,0001	0,00006	0,00006		0,00004	0,00047	0,00058		0,00002	0,00056	
	0,01707	0,00003	0,00121	0,00158	0,00298	0,00297	0,00115	0,00067	0,00155	0,0001	0,0021	0,00139	0,00504	0,00554	0,00611	0,00023	0,00096	0,00588	0,00371	0,00554	0,00011	0,00003	0,00075	0,0003	0,00091	0,00114	0,00116					0,00004	0,00019	
*	0,00006	0,001		0,00001	0,00002	0,00071	0,00387	0,00002		0,00161		0,00008	0,00003		0,00085	0,00003		0,00001	0,00022		0,00024				0,00005					0,00002				
3	0,00157	0,00449	0,00009	0,00073	0,00037	0,0007	0,00035	0,00006	0,00006	0,00039		0,00009	0,00026	0,00036	0,00179	0,00069	0,00002	0,00028		0,00001	0,00039				0,00001			0,00001	0,00047	0,00006		0,00001	0,00024	
и	0,01983	0,00011	0,00059	0,00239	0,00048	0,00151	0,0021	0,00036	0,00181	0,00122	0,00107	0,00197	0,00415	0,00298	0,00309	0,00027	0,00023	0,00127	0,00269	0,00343	0,00001	0,00002	0,00191	0,00032	0,00111	0,00052	0,00017					0,0004	0,00135	
й	0,00851					0,00028						0,00009	0,00003	0.00005	0,00046			0,00001	0,00036	0,00032		0,00001	0,00002	0,00002	0,00012	0,00005							0,00001	
к	0.00393	0.00579		0.0001	0.00001		0.00051	0.00002	0.0001	0.00283		0.00005	0.00037		0.00041	0.0078		0.00157	0.00027	0.00097	0.00139			0.00004	0.00001	0.00002								
л	0,00564	0,00387	0,00003	0,00001	0,00013	0,00003	0,00394	0,00028	0,00006	0,00698		0,0002	0,00018		0,00032	0,0051	0,00001		0,00156	0,00007	0,00123				0,00029	0,00004			0,00074	0,00362		0,00101	0,00171	
M	0,00906	0,00295	0,00006	0,00001	0,00006	0,00001	0,0036		0,00001	0,00298		0,00005	0,00018	0,00006	0,0012	0,00328	0,00014	0,00006	0,00009	0,00007	0,00254	0,00001		0,00004	0,00003				0,00098	0,00006			0,00038	
н	0.00274	0.00979			0.00001	0.00056	0.01002	0.00002	0.00025	0.00855		0.00019			0.00351	0.00954	0.00001	0.00002	0.00039	0.00084	0.00251	0.00011	0.00011	0.00115	0.00011		0.00006		0.00394	0.00134		0.00008	0.00129	
0	0.02193	0.00006	0.00353	0.00601	0.00429	0.00438	0.00217	0.00187	0.00107	0.00062	0.00376	0,00169	0.00505	0.00471	0.00544	0.0003	0.00165	0.00486	0.00613	0.00763	0.00004	0.00007	0.00064	0.0001	0.00157	0.00071	0.00033				0.0001	0.00021	0.00063	
	0.00011						0.00237			0.00053		0.00008					0.00006								0.00003	0.00001			0.00037	0.00001			0.00035	
p	0.00112	0.00707	0.00043	0.00041	0.00027	0.00023	0.00576	0.00025	0.00006	0.0043		0,00025	0.00007	0.0004	0.00119	0.00725	0.00012	0.00008	0.00023	0.00233	0.00225		0.00024	0.00003	0.00007	0.00019	0.00001		0.00148	0.00048		0.00017	0.00132	
c	0.00295	0.00129	0.00009	0.00158	0.00004	0.00016	0.00348	0.00001	0.00002	0.00131		0.00345	0.00306	0.00084	0.0014	0.00245	0.00214	0.00024	0.00067	0.01085	0.00081	0.00001	0.00012		0.00017	0.00006		0.00002	0.00028	0.00289		0.00014	0.00498	
т	0.00607	0.00712	0,00004	0.00188	0.00001	0.00008	0.00665		0.00001	0.0032		0.00072	0.00017	0.00006	0.00133	0.01343	0,00009	0.00362	0.00108	0.00006	0.00194	0.00006	0.00004	0.00002	0.00039	0.00001	0.00003		0.00183	0.00614		0,00003	0.00061	
v			0,0004					0.00137	0.00019		0,00009	0,00062	0.00145	0,0008					0.00122						0.00087						0,00002	0,00082		
ф		0.00012					0.00015			0.0002		0.00001	0.00005		0.00001	0.00043		0.00002			0.00002	0.00001							0.00001	0.00001				
×	0.00442	0.00066	0.00001	0.0001	0,00001		0.00003			0.0002			0.00013	0.00005	0,00034	0.00241		0.00017	0.00012	0.00005	0.00017					0.00001								
ц	0,00023	0.00039		0,00011			0,00052			0,00176		0,00004				0,00017					0,00006								0,00008					
	0.00028						0.00407			0.00162			0.00001	0.00001	0.00082			0.00002		0.00318						0.00011				0.00018				
	0.00008			0.00005			0.00188			0.00174					0.00049					0.00006										0.00106				
ш	0,00002	0.0006		-			0,00156			0.00098			-	,	0,00006	,	,				0,00016									0.00006				
20	.,	-,,					0.0001			.,					.,						.,									.,			0.00015	
ы	0.0045		0.00016	0.00104	0.0001	0.00011	0.00137	0.00006	0.00006		0.00147	0.00022	0.00217	0.00136	0.00011		0.00006	0.00017	0.0006	0.00071	0.00001		0.00143	0.00001	0.00022	0.00073							0.00005	
	0.01028		0.00009	-,		0,00003			0.00014		-,	0.00129	-,		0,00149	0.00003	-,	-,	0.00121		-,	0.00002	-,		0,00006		0.00001					0.00029		
,	-,-1020			0.00001		0,00000	0,5002		0,00014	-,	0.00001	0,00009	0.00007		0,00240	0,00003	0.00002		0.00003			0.00001	0.00006	0,00013	0,0000	-,	0,00001					-,	0,00017	
	0.00223		0.00019	0,00001	0.00001	0.00051		0.00009	0.00001		0,00001	0.0001	0,00007		0.00005			0.00002	0.00005				0.00002	0.00012	0.00012		0.00051					0.00001		
	0.01011			0.00029		0.00031	0.00000				0.00002	0.0001	0.00000		0.00059		0.00006			0.00174	0.00001				0.00002	0.00002						0.00009	0.00006	
-	0,01011		0,00001	0,00028	0,00015	0,00075	0,00000	0,00021	0,00028		0,00005	0,00018	0,00000	0,0006	0,00055		0,00000	0,00000	0,0006	0,00174	0,00001		0,00015	0,00002	0,00002	0,00002	0,00028					0,00000	0,00000	

Експериментуємо з програмою CoolPinkProgram.exe. Для цього проведемо по 50 експериментів для кожного випадку.  $H^{(10)}\colon$ 



# $H^{(20)}$ .





#### Висновок

У ході виконання роботи ми засвоїли поняття ентропії на символ джерела та його надлишковості, вивчили та порівняли різні моделі джерела відкритого тексту для наближеного визначення ентропії, набули практичні навички щодо оцінки ентропії на символ джерела.