



МІНІСТЕРСТВО ОСВІТИ І НАУКИ, МОЛОДІ ТА СПОРТУ УКРАЇНИ

НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ УКРАЇНИ

«КИЇВСЬКИЙ ПОЛІТЕХНІЧНИЙ ІНСТИТУТ»

ФІЗИКО-ТЕХНІЧНИЙ ІНСТИТУТ

Кафедра інформаційної безпеки

## Лабораторна робота №1

з дисципліни «Криптографія» на тему:

«Експериментальна оцінка ентропії на символ джерела  
відкритого тексту»

Перевірив:

---

Виконали:

Студентки групи ФБ-91

Легенчук М. О.

Осьмак А. А

Київ – 2021

## **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  $H(10)$ ,  $H(20)$ ,  $H(30)$ .
3. Використовуючи отримані значення ентропії, оцінити надлишковість російської мови в різних моделях джерела.

## **Хід роботи**

# Н(1)

Без пробілів			З пробілами		
Повторени			Повторени		
Буква	я	Частота	Буква	я	Частота
О	93560	0.11184	—	159100	0.15979
Е	72696	0.0869	О	93560	0.093969
А	65478	0.078271	Е	72696	0.073014
Н	56620	0.067683	А	65478	0.065764
Т	54384	0.06501	Н	56620	0.056867
И	52504	0.062762	Т	54384	0.054622
С	45924	0.054897	И	52504	0.052733
Л	37440	0.044755	С	45924	0.046125
Р	35392	0.042307	Л	37440	0.037604
В	34699	0.041479	Р	35392	0.035547
М	29485	0.035246	В	34699	0.034851
К	28387	0.033933	М	29485	0.029614
Д	25291	0.030232	К	28387	0.028511
У	24226	0.028959	Д	25291	0.025401
П	21722	0.025966	У	24226	0.024332
Я	20921	0.025009	П	21722	0.021817
Ы	16107	0.019254	Я	20921	0.021012
Ь	15692	0.018758	Ы	16107	0.016177
Г	15186	0.018153	Ь	15692	0.015761
Б	14017	0.016756	Г	15186	0.015252
З	13705	0.016383	Б	14017	0.014078
Ч	13108	0.015669	З	13705	0.013765
Й	10040	0.012002	Ч	13108	0.013165
Ж	8684	0.010381	Й	10040	0.010084
		0.009126			0.008721
Х	7635	8	Ж	8684	9
		0.008529			0.007668
Ш	7135	1	Х	7635	3
		0.007649			0.007166
Ю	6399	3	Ш	7135	2
		0.003839			
Э	3212	6	Ю	6399	0.006427
		0.003447			
Ц	2884	5	Э	3212	0.003226
		0.003146			0.002896
Щ	2632	3	Ц	2884	6
		0.001656			0.002643
Ф	1386	8	Щ	2632	5
					0.001392
			Ф	1386	1
Энтропия:			Энтропия:		
4.4695			4.3891		



З проблемами та кроком																																	
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	Ъ	Ы	Ь	Э	Ю	Я		
3.5858e-05	0.0019258	0.006001	0.014729	0.0036828	0.0070915	0.0036575	0.0015862	0.0046742	0.0088464	4.2186e-06	0.0089455	0.002567	0.0091902	0.015666	0.01177	0.015092	0.0037946	0.015231	0.006813	0.0049315	0.00070029	0.0022802	0.00032694	0.0051636	0.0010293	9.7028e-05	-	-	-	0.0025396	6.3279e-05	0.004805	
А	0.011711	-	0.00084794	0.0025965	0.00079732	0.002219	0.0028539	0.0010884	0.0032462	0.00017507	0.00082685	0.004554	0.006342	0.002972	0.0044844	2.1093e-05	0.00087536	0.0042988	0.0041532	0.00010968	0.00014132	0.00092387	0.00012656	0.00091544	0.0011285	0.00026999	-	-	-	-	0.0022949	0.001814	
Б	0.00017929	0.00075302	9.2809e-05	4.2186e-05	-	2.953e-05	0.0022169	1.2656e-05	2.1093e-06	0.00083528	-	0.00019195	0.00085638	7.8044e-05	0.00035436	0.0024932	-	0.0011411	0.0001582	1.0546e-05	0.0011264	-	5.2732e-05	4.2186e-06	8.4372e-06	1.2656e-05	0.00017507	0.003145	0.00015609	2.1093e-06	4.2186e-06	0.00051256	
В	0.004284	0.0060769	8.4372e-06	3.3749e-05	6.3279e-06	0.0002257	0.0047037	-	0.00051467	0.0029319	-	0.00025733	0.0004497	0.00012445	0.00059404	0.00024468	0.00062857	0.00025101	0.00068763	-	6.5388e-05	2.3202e-05	4.8514e-05	0.00049358	2.1093e-06	0.0002046	-	-	-	-	0.0002953		
Г	0.00047459	0.0011791	-	8.4372e-06	1.8984e-05	0.0012255	0.00022991	-	4.2186e-06	0.00083528	-	0.00014976	0.0011791	1.2656e-05	0.00020429	0.0071646	-	0.0013078	5.2732e-05	2.953e-05	0.00069607	-	8.4372e-06	-	4.8514e-05	-	-	-	-	-	-		
Д	0.00078044	0.0042671	1.6405e-05	0.00092387	6.3279e-06	5.906e-05	0.0048408	0.00064756	1.6874e-05	0.0025185	-	0.00026366	0.00070662	4.0077e-05	0.0017296	0.0040667	0.00013289	0.0011243	0.00033749	0.00012023	0.0015524	6.3279e-06	5.0623e-05	0.00024468	5.0623e-05	6.7498e-05	-	-	0.00064756	0.00054842	-	3.1639e-05	0.00034171
Е	0.014508	7.8044e-05	0.0012593	0.0013162	0.0036955	0.002837	0.001139	0.00071083	0.0012825	9.07e-05	0.0027442	0.001177	0.0053618	0.0030971	0.0079267	0.00025733	0.00092809	0.00069438	0.0043198	0.00074732	0.00012867	0.00011179	0.00054631	0.00039444	0.0010947	0.00087958	0.00068763	-	-	-	0.00021515	0.00021726	
Ж	0.000135	0.0011707	6.3279e-06	4.2186e-06	1.0546e-05	0.00083528	0.0034171	1.6874e-05	-	0.0015039	-	0.00014765	2.7421e-05	6.3279e-06	0.00018351	2.1093e-06	2.1093e-06	4.2186e-06	-	0.0004535	-	-	5.6951e-05	-	-	-	-	-	-	-	-	-	
З	0.003162	0.005442	0.00014554	0.00067989	0.00037756	0.00074669	0.00042186	0.00010336	1.2656e-05	0.00032061	-	0.00071718	0.00027421	0.00036491	0.0017001	0.00054842	-	0.00024257	6.3279e-06	1.4765e-05	0.00041131	-	4.2186e-06	1.4765e-05	2.1093e-06	-	-	0.00040499	0.00014132	-	0.00020671		
И	0.011397	0.00079521	0.00050623	0.0025944	0.00050834	0.0016242	0.00034382	0.0020819	0.00057162	0.0010504	-	0.0019511	0.003936	0.00016874	0.00019406	0.00070451	0.00073335	0.0047712	2.7421e-05	9.2809e-05	0.0015039	0.00094708	0.00099348	0.00055475	0.00025312	-	-	-	-	4.8514e-05	0.00037335	0.001297	
Й	0.0058301	-	6.3279e-06	2.1093e-06	-	0.00027843	8.4372e-06	-	2.3202e-05	4.2186e-06	-	0.00012656	0.001371	9.3137e-05	0.00038178	0.00010336	2.1093e-05	-	0.00051678	0.00053365	-	2.1093e-06	-	7.1716e-05	0.0002046	0.00027843	-	-	-	-	-	-	
К	0.0035942	0.0027602	1.8984e-05	0.00020882	-	0.00076357	3.3749e-05	3.5858e-05	0.0024911	-	5.4842e-05	0.00080364	6.3279e-06	0.00045983	0.0078846	-	0.0025902	0.00015398	0.00066601	0.0017781	2.1093e-06	2.3202e-05	3.7967e-05	-	1.0546e-05	-	-	-	-	4.2186e-06	0.00010546	2.1093e-06	-
Л	0.0045666	0.0049885	3.7967e-05	1.0546e-05	0.0012023	6.3279e-05	0.0043241	0.00028476	2.7421e-05	0.0061887	-	0.00038389	0.00025101	0.00059482	0.00035225	0.005731	2.953e-05	-	0.0011285	0.00016031	0.0015503	4.2186e-06	2.1093e-06	-	8.0153e-05	4.2186e-06	2.1093e-06	0.0009745	0.0039317	4.4295e-05	0.00093231	0.0015609	
М	0.0054525	0.0028539	3.7967e-05	-	4.6405e-05	-	0.0046384	-	8.4372e-06	0.0030079	-	0.00013289	0.00020308	3.7825e-05	0.0024911	0.004748	0.00086692	1.6874e-05	0.0020777	2.7421e-05	2.1093e-06	1.8984e-05	7.1716e-05	6.3279e-06	1.0546e-05	0.0014575	8.8591e-05	1.4765e-05	2.1093e-06	0.00039655	-	-	
Н	0.0040035	0.011291	6.3279e-06	5.6951e-05	0.00021937	0.00061381	0.010158	1.4765e-05	1.6874e-05	0.0068468	-	0.00038178	2.1093e-06	-	0.0024974	0.010308	4.2186e-06	0.00012445	0.00085638	0.0025058	1.2656e-05	1.4765e-05	0.00027421	0.00020882	1.2656e-05	9.07e-05	0.0034128	0.0010905	-	-	0.00017085	0.0025291	
О	0.018321	8.4372e-06	0.0036554	0.0058048	0.004497	0.0046162	0.0024573	0.0021536	0.0013647	0.0010968	0.0038917	0.00023835	0.0062709	0.0052311	0.0072497	0.00022991	0.0011432	0.0050918	0.0058428	0.0062744	0.00020671	0.00017507	0.00055053	0.00012023	0.0023118	0.00090278	0.00023835	-	-	7.5935e-05	0.00037967	0.00074247	
П	5.0623e-05	0.0015567	-	-	-	0.0018035	-	-	0.00095551	-	0.00013289	0.00094708	-	0.00017929	0.0093611	3.1639e-05	0.0064249	6.3279e-06	5.2732e-05	0.00676779	2.1093e-06	-	2.1093e-06	9.7028e-05	6.3279e-06	-	-	-	-	0.00035858	8.2263e-05	2.1093e-06	0.00029952
Р	0.00078677	0.0073678	5.2732e-05	0.00035225	0.0002257	0.00040499	0.00036491	5.0623e-05	0.0049653	-	0.00030585	0.00029741	0.00037335	0.00085427	0.00068236	0.00015398	8.4372e-05	0.00039655	0.00085427	0.0026155	2.1093e-05	0.00014132	3.1639e-05	8.4372e-05	0.0002148	1.4765e-05	0.0013162	0.00063279	7.5935e-05	0.00024257	0.00096817	-	
С	0.002856	0.0015904	7.1716e-05	0.0014238	2.3202e-05	0.0002489	0.0034698	6.5388e-05	2.1093e-06	0.001736	-	0.00035563	0.0028792	0.0011348	0.0012255	0.0028096	0.001736	0.00024679	0.0010399	0.012	0.00069818	2.3202e-05	0.00017507	4.6405e-05	0.00033749	9.7028e-05	-	-	0.00050201	0.002991	0.00013139	0.00016453	0.0040836
Т	0.0064587	0.0056023	5.0623e-05	0.0022865	1.4765e-05	0.00012023	0.0062667	-	6.3279e-06	0.0041996	-	0.00066443	0.00036069	4.8514e-05	0.0013162	0.012696	0.00021656	0.0030437	0.0021768	8.6481e-05	0.0013774	8.4372e-06	8.4372e-06	0.00010968	0.0001139	1.2656e-05	4.4295e-05	0.0007739	0.0058027	0.0001139	5.0623e-05	0.00063279	
У	0.0049062	6.5388e-05	0.00083106	0.00073615	0.00097661	0.0019005	0.00036913	0.0013352	0.00036491	0.00010757	0.00024046	0.00093653	0.001371	0.0010483	0.00025101	6.7498e-05	0.00080364	0.00048514	0.0012529	0.0014807	6.3279e-06	1.6874e-05	0.00034803	2.1093e-06	0.0008627	0.00060537	0.00028265	-	-	4.8514e-05	0.0010399	4.4295e-05	
Ф	1.6874e-05	0.00051587	-	-	1.8984e-05	-	0.00027843	-	-	0.00024046	-	-	4.8514e-05	2.1093e-06	2.1093e-06	0.00034382	-	8.2263e-05	2.1093e-06	4.8514e-05	8.0153e-05	-	4.2186e-06	-	-	-	-	-	1.6874e-05	2.1093e-06	-	2.1093e-06	
Х	0.0022738	0.00058217	-	0.00021726	2.1093e-06	-	3.3749e-05	-	-	0.00022991	-	-	9.2809e-05	0.00018984	0.00017296	0.0031492	-	0.00012867	3.1639e-05	1.2656e-05	0.00012445	-	6.3279e-06	-	2.1093e-06	1.6874e-05	-	-	-	4.2186e-06	-	-	
Ц	0.0002038	0.0061381	-	7.8044e-05	-	2.1093e-06	0.00087536	-	-	0.00039866	-	1.8984e-05	-	-	2.1093e-06	-	0.00026366	2.1093e-06	-	-	-	-	-	-	-	-	-	-	-	0.00021093	-	-	
Ч	2.7421e-05	0.0021072	-	4.2186e-06	-	-	0.0041743	-	-	0.0015967	-	0.00031639	1.6874e-05	2.1093e-06	0.0009534	6.7498e-05	-	8.4372e-06	-	0.0034698	0.00083528	-	-	-	-	-	-	-	0.00013289	-	-	0.00026788	-
Ш	0.00015398	0.0010399	-	1.6874e-05	-	-	0.0022591	-	-	0.0013879	-	0.00041131	0.00049358	1.6874e-05	0.00031639	0.0004535	1.0546e-05	-	5.6951e-05	0.00031007	-	1.4765e-05	-	-	-	-	-	-	-	0.00054842	-	4.2186e-06	-
Щ	4.2186e-06	0.00043663	-	-	-	0.0012171	-	-	0.00078044	-	-	-	-	-	6.6481e-05	2.1093e-06	-	-	-	-	-	-	-	-	-	-	-	-	-	-	6.3279e-05	-	-
Ъ	0.004478	-	0.00039444	0.0011686	0.00014976	0.00010336	0.0010779	6.7498e-05	6.117e-05	6.3279e-06	0.001563	0.00022991	0.0016389	0.0010715	0.00018351	-	0.0001371	0.00027421	0.00064966	0.00067709	4.2186e-06	-	0.00098504	8.4372e-06	0.00021515	0.00043452	4.4295e-05	-	-	-	-	3.1639e-05	
Ы	0.0072201	-	7.1716e-05	1.8984e-05	3.7967e-05	4.6405e-05	0.0004999	-	0.00013921	0.4077e-05	-	0.00094708	-	0.00029319	0.0014217	1.6874e-05	-	2.1093e-06	-	0.00097239	0.0001135	-	6.3279e-06	-	0.00020671	5.6951e-05	0.00050623	1.0546e-05	-	-	0.00048936	0.00035858	-
Ь	2.186e-06	-	9.7028e-05	2.1093e-06	1.6874e-05	1.0546e-05	2.1093e-06	-	1.4765e-05	-	-	8.4514e-05	0.00011812	0.00010546	0.00077e-05	-	1.2656e-05	0.00013289	2.7421e-05	0.0024679	-	-	-	-	-	-	-	-	-	2.1093e-06	6.3279e-06	-	
Э	0.0027273	-	0.00030163	4.2186e-06	2.1093e-05	0.0003628																											

	Текст з пробілами	Текст без пробілів
<b>H1</b>	4.3891	4.4695
<b>H2</b>	4.0201	3.9927
<b>H2 (з кроком 2)</b>	4.0196	3.9922

CoolPinkProgram

H(10): 1.7835 < H < 2.5509

Лабораторная работа №1

Произвольная часть текста:  
м\_внешними\_причинами\_мы\_ставим\_его\_исключительно\_в\_заслугу\_себе\_итак\_я\_хочу

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: и

Символ по счету: 2

Номер эксперимента: 50

Поле ввода символов:  
и

Продолжить Другой

Неравенство для энтропии:  
1,78357968437524< H < 2,55092368847664

Двоичная таблица угаданных символов:  
10000000000000000000000000000000  
10000000000000000000000000000000  
00000001000000000000000000000000  
10000000000000000000000000000000  
00000000100000000000000000000000

Вероятности:  
q[ 1 ] = 0,54  
q[ 2 ] = 0,08  
q[ 3 ] = 0,06  
q[ 4 ] = 0,06  
q[ 5 ] = 0,06  
q[ 6 ] = 0  
q[ 7 ] = 0  
q[ 8 ] = 0,02  
q[ 9 ] = 0,02  
q[ 10 ] = 0  
q[ 11 ] = 0,04  
q[ 12 ] = 0  
q[ 13 ] = 0  
q[ 14 ] = 0,02  
q[ 15 ] = 0  
q[ 16 ] = 0  
q[ 17 ] = 0  
q[ 18 ] = 0,04  
q[ 19 ] = 0  
q[ 20 ] = 0,02  
q[ 21 ] = 0,02  
q[ 22 ] = 0  
q[ 23 ] = 0,02  
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

[illegible]

**Лабораторная работа №1**

---

Произвольная часть текста:  
**ду\_справедливыми\_и\_несправедливыми\_договорами\_я\_думаю\_шила\_в\_мешке\_не\_утаиш**

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: **и**

Символ по счету: **1**

Номер эксперимента: **50**

Неравенство для энтропии:  
**1,99867595623896 < H < 2,62262668981542**

Двоичная таблица угаданных символов:

10000000000000000000000000000000	▲
00100000000000000000000000000000	
00000000000000000001000000000000	
10000000000000000000000000000000	
01000000000000000000000000000000	▼

Вероятности:

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

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

## Оцінка надлишковості:

	Текст з пробілами, R	Текст без пробілів, R
<b>H1</b>	0.12218	0.1061
<b>H2</b>	0.19598	0.20146
<b>H2 (з кроком 2)</b>	0.19608	0.20156

## Висновки:

Для отриманих значень було використано алфавіт російської мови в якому замінені букви «ъ» «ё» та врахований пробіл. Довжина вхідного алфавіту 32 символу.

Для даної мови (російської) розраховано, що надлишковість встановлює 0.17056.

Встановлена залежність: при збільшенні  $n$  питомої ентропії її значення зменшується. Так аналітично можливо визначити наступні члени, оскільки функція є неперервною.