

**«Криптографія»
Комп'ютерний практикум №1**

Експериментальна оцінка ентропії на символ джерела відкритого тексту

Виконали:
студенти групи ФБ-01
Корабельський Тарас
та
Оліферчук Владислав
варіант №10

Мета роботи

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

Постановка задачі:

Визначити надлишковість російської мови в різних моделях джерела, вираховуючи за допомогою самостійно написаної програми частоту входження буквенних символів в текст (розмір від 1 Мб, визначення ентропію H_1 та H_2 для букв та біграм (з букв, які перетинаються, і не перетинаються) з урахуванням пробілів і без них, за допомогою спеціальної програми CoolPinkProgram оцінити H^{10}, H^{20}, H^{30}

Для лабораторної роботи використовувався текст:
«Шинель» М.В. Гоголя.

Джерело	H	R
H1 текст з пробілами	4.355	0.129
H1 текст без пробілів	4.433	0.113
H2 перехресна біграма з пробілами	3.8984	0.2203
H2 не перехресна біграма з пробілами	3.9137	0.2172
H2 перехресна біграма без пробілів	4.0693	0.1861
H2 не перехресна біграма без пробілів	4.0613	0.1877

Ч

Частоти літер у тексті з пробілами:

{ ' ': 0.16079456684628385,
'о': 0.09549469119663298,
'е': 0.07374932245002072,
'а': 0.06722890029652775,
'и': 0.059831648758090744,
'н': 0.055319963013742306,
'т': 0.050027101999171,
'с': 0.041163154035009406,
'л': 0.04015878583043714,
'в': 0.03819787647865319,
'к': 0.03791091413448969,

'р': 0.03249051430029015,
'м': 0.025364282753563117,
'д': 0.025093262761853138,
'у': 0.024009182795013233,
'п': 0.021888849918694,
'ч': 0.017871377100404935,
'ь': 0.017568472403787904,
'г': 0.014539425437617574,
'я': 0.01449159838025699,
'б': 0.013822019577208813,
'з': 0.01288142078245066,
'ж': 0.009342218537767432,
'ш': 0.008608870324905144,
'й': 0.008592927972451614,
'х': 0.006376940981411217,
'ю': 0.004097184580556707,
'ц': 0.0032841246054267767,
'щ': 0.0024072952204827343,
'ф': 0.00047827057360584125}
H1 = 4.252

Частоти літер у тексті без пробілів:

{'о': 0.1137853085880364,
'е': 0.08787493113994263,
'а': 0.0801056170810934,
'и': 0.07129152973804684,
'н': 0.06591569629390422,
'т': 0.059609064832931255,
'с': 0.04904735672359098,
'л': 0.047850616416237675,
'в': 0.045514123435214555,
'к': 0.045172197633113614,
'р': 0.03871359914898467,
'м': 0.03022244173014456,
'д': 0.02989951180593811,
'у': 0.028607792109112323,
'п': 0.026081340349144236,
'ч': 0.02129437911973102,
'ь': 0.020933457439735577,
'г': 0.017324240639781167,
'я': 0.017267253006097678,
'б': 0.016469426134528807,
'з': 0.015348669338753491,
'ж': 0.01113158444617518,

'ш': 0.010257774063028323,
'й': 0.010238778185133825,
'х': 0.007598351157798757,
'ю': 0.004881940618885702,
'ц': 0.003913150846266361,
'щ': 0.002868377562069031,
'ф': 0.0005698763368349068}
H1 = 4.315

Частоти перехресних біграм у тексті з пробілами

/	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	ь	э	ю	я			
а	0.0000	0.0007	0.0025	0.0006	0.0015	0.0009	0.0019	0.0037	0.0000	0.0005	0.0130	0.0072	0.0038	0.0025	0.0000	0.0007	0.0024	0.0026	0.0045	0.0000	0.0000	0.0010	0.0000	0.0017	0.0007	0.0003	0.0000	0.0000	0.0000	0.0005	0.0014	0.0132	
б	0.0006	0.0000	0.0000	0.0000	0.0000	0.0017	0.0000	0.0000	0.0004	0.0000	0.0001	0.0007	0.0000	0.0003	0.0017	0.0000	0.0013	0.0000	0.0019	0.0000	0.0000	0.0000	0.0000	0.0001	0.0042	0.0000	0.0000	0.0000	0.0000	0.0002	0.0003	0.0000	
в	0.0051	0.0000	0.0000	0.0000	0.0003	0.0057	0.0000	0.0004	0.0050	0.0000	0.0002	0.0004	0.0002	0.0014	0.0057	0.0003	0.0007	0.0033	0.0001	0.0006	0.0000	0.0000	0.0000	0.0012	0.0000	0.0020	0.0002	0.0000	0.0000	0.0002	0.0052	0.0051	
г	0.0005	0.0000	0.0000	0.0000	0.0012	0.0003	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001	0.0006	0.0000	0.0005	0.0001	0.0000	0.0004	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0011	
д	0.0050	0.0000	0.0007	0.0000	0.0000	0.0053	0.0000	0.0000	0.0020	0.0000	0.0004	0.0007	0.0001	0.0019	0.0032	0.0000	0.0013	0.0003	0.0004	0.0012	0.0000	0.0000	0.0001	0.0001	0.0000	0.0000	0.0004	0.0007	0.0000	0.0003	0.0009	0.0000	
е	0.0000	0.0010	0.0023	0.0035	0.0024	0.0014	0.0008	0.0010	0.0002	0.0015	0.0014	0.0000	0.0047	0.0070	0.0002	0.0012	0.0051	0.0044	0.0052	0.0001	0.0000	0.0006	0.0007	0.0013	0.0006	0.0000	0.0000	0.0000	0.0000	0.0002	0.0011	0.0178	
ж	0.0010	0.0000	0.0000	0.0000	0.0006	0.0005	0.0000	0.0000	0.0013	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0006	0.0006	
з	0.0045	0.0001	0.0000	0.0002	0.0007	0.0003	0.0000	0.0000	0.0003	0.0000	0.0001	0.0002	0.0001	0.0020	0.0006	0.0000	0.0001	0.0000	0.0000	0.0002	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0006	0.0014
и	0.0000	0.0007	0.0022	0.0004	0.0016	0.0030	0.0002	0.0020	0.0004	0.0018	0.0025	0.0048	0.0028	0.0044	0.0000	0.0001	0.0006	0.0022	0.0035	0.0000	0.0000	0.0015	0.0017	0.0033	0.0004	0.0002	0.0000	0.0000	0.0000	0.0004	0.0017	0.0174	
й	0.0000	0.0000	0.0000	0.0000	0.0001	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0003	0.0004	0.0000	0.0000	0.0000	0.0001	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0071	
к	0.0115	0.0000	0.0003	0.0000	0.0000	0.0005	0.0000	0.0000	0.0059	0.0000	0.0000	0.0004	0.0000	0.0000	0.0000	0.0000	0.0000	0.0014	0.0001	0.0000	0.0018	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0054	
л	0.0040	0.0000	0.0000	0.0002	0.0000	0.0033	0.0003	0.0001	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001	0.0006	0.0000	0.0000	0.0019	0.0000	0.0014	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0006	0.0014	0.0070	
м	0.0010	0.0000	0.0000	0.0000	0.0000	0.0038	0.0000	0.0000	0.0022	0.0000	0.0000	0.0000	0.0000	0.0001	0.0011	0.0030	0.0001	0.0000	0.0000	0.0020	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0005	0.0087
н	0.0093	0.0000	0.0000	0.0000	0.0002	0.0104	0.0000	0.0000	0.0077	0.0000	0.0000	0.0000	0.0000	0.0031	0.0115	0.0000	0.0000	0.0002	0.0005	0.0021	0.0000	0.0000	0.0002	0.0001	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
о	0.0000	0.0033	0.0085	0.0050	0.0042	0.0020	0.0017	0.0008	0.0009	0.0027	0.0015	0.0054	0.0061	0.0054	0.0001	0.0009	0.0051	0.0053	0.0069	0.0001	0.0000	0.0003	0.0000	0.0021	0.0011	0.0003	0.0000	0.0000	0.0000	0.0007	0.0055	0.0245	
п	0.0010	0.0000	0.0000	0.0000	0.0000	0.0028	0.0000	0.0000	0.0011	0.0000	0.0001	0.0007	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	
р	0.0057	0.0002	0.0002	0.0002	0.0003	0.0057	0.0002	0.0000	0.0041	0.0000	0.0002	0.0001	0.0000	0.0006	0.0006	0.0000	0.0000	0.0002	0.0012	0.0028	0.0000	0.0001	0.0000	0.0001	0.0006	0.0000	0.0013	0.0004	0.0000	0.0001	0.0009	0.0000	
с	0.0015	0.0000	0.0015	0.0000	0.0005	0.0031	0.0000	0.0000	0.0012	0.0000	0.0011	0.0028	0.0010	0.0007	0.0025	0.0014	0.0002	0.0005	0.0097	0.0008	0.0002	0.0000	0.0002	0.0001	0.0000	0.0000	0.0003	0.0030	0.0000	0.0001	0.0044	0.0022	
т	0.0060	0.0000	0.0022	0.0000	0.0001	0.0044	0.0000	0.0000	0.0028	0.0000	0.0003	0.0002	0.0000	0.0017	0.0146	0.0001	0.0032	0.0010	0.0001	0.0013	0.0000	0.0000	0.0000	0.0001	0.0000	0.0000	0.0000	0.0000	0.0000	0.0004	0.0044	0.0044	
у	0.0000	0.0006	0.0008	0.0013	0.0021	0.0002	0.0021	0.0003	0.0000	0.0000	0.0001	0.0007	0.0015	0.0011	0.0002	0.0000	0.0002	0.0004	0.0011	0.0012	0.0000	0.0000	0.0007	0.0000	0.0010	0.0004	0.0001	0.0000	0.0000	0.0000	0.0011	0.0005	
ф	0.0001	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	
х	0.0005	0.0000	0.0002	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001	0.0000	0.0000	0.0000	0.0002	0.0020	0.0000	0.0000	0.0000	0.0000	0.0001	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	
ц	0.0007	0.0000	0.0000	0.0000	0.0000	0.0008	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0004	0.0000	0.0000	0.0000	0.0000	0.0004	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	
ч	0.0020	0.0000	0.0000	0.0000	0.0000	0.0036	0.0000	0.0000	0.0026	0.0000	0.0004	0.0000	0.0000	0.0010	0.0001	0.0000	0.0001	0.0000	0.0042	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0015	
ш	0.0000	0.0000	0.0001	0.0000	0.0000	0.0020	0.0000	0.0000	0.0029	0.0000	0.0004	0.0005	0.0001	0.0003	0.0002	0.0000	0.0000	0.0000	0.0000	0.0000	0.0003	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	
щ	0.0002	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0006	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	
ь	0.0000	0.0001	0.0011	0.0001	0.0001	0.0010	0.0000	0.0001	0.0000	0.0017	0.0004	0.0022	0.0015	0.0002	0.0000	0.0002	0.0002	0.0005	0.0005	0.0000	0.0000	0.0007	0.0000	0.0002	0.0005	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0031
э	0.0000	0.0001	0.0000	0.0000	0.0000	0.0001	0.0001	0.0003	0.0000	0.0002	0.0000	0.0001	0.0000	0.0001	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
ю	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
я	0.0000	0.0002	0.0000	0.0000	0.0000	0.0002	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0028	0.0028
а	0.0000	0.0000	0.0003	0.0000	0.0004	0.0001	0.0001	0.0000	0.0000	0.0000	0.0001	0.0005	0.0006	0.0004	0.0005	0.0000	0.0000	0.0003	0.0004	0.0015	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
б	0.0017	0.0006	0.0144	0.0029	0.0005	0.0043	0.0019	0.0039	0.0104	0.0000	0.0099	0.0026	0.0038	0.0163	0.0100	0.0162	0.0034	0.0154	0.0071	0.0040	0.0003	0.0012	0.0003	0.0074	0.0019	0.0001	0.0000	0.0000	0.0010	0.0000	0.0005	0.0015	0.0015
H2= 3.8984																																	

Частоти не перехресних біграм у тексті з пробілами

/	а	б	в	г
---	---	---	---	---

Частоти не перехресних біграм у тексті без пробілів

	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	0.0002	0.0019	0.0046	0.0008	0.0027	0.0018	0.0026	0.0044	0.0011	0.0005	0.0163	0.0088	0.0037	0.0047	0.0008	0.0025	0.0032	0.0042	0.0067	0.0003	0.0000	0.0013	0.0001	0.0026	0.0010	0.0004	0.0000	0.0000	0.0002	0.0003	0.0016
б	0.0009	0.0001	0.0000	0.0000	0.0000	0.0019	0.0000	0.0000	0.0006	0.0000	0.0000	0.0000	0.0010	0.0000	0.0003	0.0023	0.0000	0.0015	0.0003	0.0000	0.0022	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0002	0.0049	0.0000	0.0001
в	0.0005	0.0000	0.0003	0.0004	0.0009	0.0070	0.0001	0.0005	0.0057	0.0000	0.0008	0.0008	0.0003	0.0019	0.0076	0.0009	0.0010	0.0041	0.0008	0.0009	0.0000	0.0000	0.0000	0.0003	0.0015	0.0000	0.0025	0.0003	0.0002	0.0000	0.0001
г	0.0006	0.0001	0.0002	0.0000	0.0014	0.0004	0.0000	0.0001	0.0012	0.0000	0.0001	0.0010	0.0000	0.0003	0.0101	0.0002	0.0006	0.0003	0.0000	0.0005	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001
д	0.0050	0.0000	0.0008	0.0001	0.0001	0.0054	0.0000	0.0002	0.0024	0.0000	0.0005	0.0010	0.0000	0.0024	0.0035	0.0002	0.0015	0.0003	0.0005	0.0016	0.0000	0.0000	0.0002	0.0001	0.0001	0.0000	0.0005	0.0010	0.0000	0.0000	0.0002
е	0.0002	0.0024	0.0042	0.0044	0.0041	0.0024	0.0013	0.0022	0.0017	0.0017	0.0029	0.0109	0.0062	0.0098	0.0017	0.0035	0.0063	0.0075	0.0070	0.0006	0.0001	0.0009	0.0007	0.0025	0.0009	0.0009	0.0000	0.0000	0.0002	0.0003	0.0002
ж	0.0011	0.0000	0.0001	0.0000	0.0000	0.0056	0.0000	0.0000	0.0015	0.0000	0.0003	0.0000	0.0000	0.0010	0.0001	0.0000	0.0000	0.0000	0.0000	0.0001	0.0000	0.0000	0.0000	0.0001	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
з	0.0051	0.0003	0.0011	0.0003	0.0009	0.0002	0.0000	0.0001	0.0004	0.0000	0.0003	0.0002	0.0000	0.0027	0.0008	0.0001	0.0001	0.0001	0.0003	0.0003	0.0000	0.0000	0.0000	0.0000	0.0000	0.0002	0.0000	0.0000	0.0000	0.0000	0.0000
и	0.0003	0.0014	0.0050	0.0000	0.0034	0.0040	0.0003	0.0030	0.0016	0.0023	0.0043	0.0059	0.0030	0.0070	0.0013	0.0020	0.0014	0.0044	0.0050	0.0005	0.0002	0.0020	0.0023	0.0046	0.0005	0.0002	0.0000	0.0000	0.0004	0.0005	0.0020
й	0.0013	0.0002	0.0005	0.0003	0.0006	0.0001	0.0002	0.0001	0.0007	0.0000	0.0005	0.0001	0.0005	0.0010	0.0001	0.0008	0.0002	0.0013	0.0007	0.0001	0.0000	0.0000	0.0001	0.0005	0.0005	0.0000	0.0000	0.0000	0.0001	0.0000	0.0001
к	0.0145	0.0005	0.0007	0.0001	0.0000	0.0006	0.0006	0.0002	0.0002	0.0077	0.0000	0.0004	0.0004	0.0002	0.0020	0.0111	0.0005	0.0018	0.0011	0.0011	0.0020	0.0000	0.0001	0.0000	0.0005	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
л	0.0051	0.0003	0.0009	0.0005	0.0003	0.0041	0.0004	0.0003	0.0001	0.0000	0.0010	0.0002	0.0000	0.0012	0.0005	0.0009	0.0001	0.0025	0.0005	0.0016	0.0000	0.0001	0.0000	0.0000	0.0000	0.0001	0.0000	0.0007	0.0000	0.0007	0.0015
м	0.0025	0.0003	0.0008	0.0002	0.0011	0.0045	0.0002	0.0003	0.0036	0.0000	0.0000	0.0001	0.0002	0.0020	0.0050	0.0011	0.0002	0.0014	0.0003	0.0036	0.0000	0.0000	0.0001	0.0003	0.0000	0.0000	0.0000	0.0000	0.0002	0.0000	0.0006
н	0.0109	0.0005	0.0005	0.0001	0.0003	0.0129	0.0000	0.0002	0.0100	0.0000	0.0005	0.0000	0.0000	0.0040	0.0141	0.0006	0.0001	0.0007	0.0008	0.0026	0.0000	0.0001	0.0003	0.0002	0.0001	0.0001	0.0040	0.0019	0.0000	0.0002	0.0009
о	0.0005	0.0052	0.0133	0.0066	0.0066	0.0036	0.0023	0.0022	0.0023	0.0032	0.0033	0.0058	0.0076	0.0088	0.0023	0.0037	0.0069	0.0087	0.0092	0.0011	0.0000	0.0007	0.0000	0.0041	0.0014	0.0002	0.0000	0.0000	0.0003	0.0010	0.0006
п	0.0016	0.0000	0.0000	0.0000	0.0000	0.0033	0.0000	0.0000	0.0011	0.0000	0.0002	0.0010	0.0000	0.0001	0.0112	0.0000	0.0070	0.0000	0.0000	0.0008	0.0000	0.0000	0.0001	0.0000	0.0000	0.0000	0.0001	0.0000	0.0000	0.0000	0.0004
р	0.0006	0.0003	0.0004	0.0003	0.0003	0.0006	0.0003	0.0000	0.0052	0.0000	0.0005	0.0002	0.0001	0.0011	0.0076	0.0001	0.0001	0.0002	0.0014	0.0040	0.0000	0.0001	0.0000	0.0001	0.0000	0.0000	0.0015	0.0006	0.0000	0.0001	0.0013
с	0.0021	0.0000	0.0021	0.0001	0.0000	0.0030	0.0001	0.0002	0.0013	0.0000	0.0043	0.0039	0.0013	0.0014	0.0030	0.0017	0.0003	0.0006	0.0117	0.0011	0.0000	0.0002	0.0000	0.0003	0.0000	0.0000	0.0005	0.0035	0.0001	0.0001	0.0050
т	0.0071	0.0002	0.0030	0.0000	0.0003	0.0054	0.0002	0.0002	0.0036	0.0000	0.0010	0.0004	0.0001	0.0023	0.0188	0.0006	0.0036	0.0020	0.0000	0.0013	0.0000	0.0000	0.0000	0.0003	0.0000	0.0000	0.0011	0.0009	0.0001	0.0000	0.0005
у	0.0002	0.0008	0.0018	0.0016	0.0030	0.0003	0.0028	0.0005	0.0006	0.0001	0.0012	0.0019	0.0015	0.0015	0.0004	0.0016	0.0005	0.0021	0.0017	0.0001	0.0000	0.0007	0.0001	0.0022	0.0003	0.0002	0.0000	0.0000	0.0000	0.0001	0.0001
ф	0.0002	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001	0.0000	0.0000	0.0000	0.0001	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
х	0.0006	0.0002	0.0005	0.0001	0.0000	0.0002	0.0000	0.0000	0.0005	0.0000	0.0001	0.0000	0.0002	0.0001	0.0000	0.0025	0.0007	0.0001	0.0004	0.0002	0.0001	0.0000	0.0000	0.0002	0.0000	0.0000	0.0000	0.0000	0.0001	0.0000	0.0000
ц	0.0008	0.0000	0.0002	0.0000	0.0000	0.0009	0.0000	0.0000	0.0002	0.0000	0.0001	0.0000	0.0002	0.0001	0.0005	0.0001	0.0000	0.0000	0.0000	0.0004	0.0000	0.0000	0.0000	0.0001	0.0000	0.0000	0.0000	0.0001	0.0000	0.0000	0.0001
ч	0.0032	0.0000	0.0001	0.0000	0.0001	0.0047	0.0000	0.0000	0.0000	0.0030	0.0000	0.0005	0.0000	0.0000	0.0014	0.0002	0.0003	0.0004	0.0001	0.0050	0.0011	0.0000	0.0000	0.0000	0.0000	0.0000	0.0003	0.0000	0.0000	0.0000	0.0000
ш	0.0008	0.0000	0.0002	0.0000	0.0000	0.0038	0.0000	0.0000	0.0035	0.0000	0.0005	0.0005	0.0002	0.0003	0.0002	0.0000	0.0000	0.0000	0.0000	0.0001	0.0004	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
щ	0.0003	0.0000	0.0000	0.0000	0.0000	0.0017	0.0000	0.0000	0.0005	0.0000	0.0000	0.0000	0.0000	0.0001	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
ы	0.0001	0.0002	0.0015	0.0001	0.0003	0.0010	0.0000	0.0002	0.0004	0.0020	0.0000	0.0020	0.0018	0.0006	0.0002	0.0004	0.0003	0.0012	0.0000	0.0000	0.0000	0.0000	0.0000	0.0003	0.0006	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
ь	0.0002	0.0005	0.0010	0.0002	0.0011	0.0011	0.0001	0.0006	0.0011	0.0000	0.0021	0.0003	0.0007	0.0033	0.0011	0.0009	0.0003	0.0022	0.0007	0.0005	0.0000	0.0001	0.0001	0.0011	0.0000	0.0000	0.0000	0.0000	0.0001	0.0002	0.0002
э	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0016	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
ю	0.0003	0.0003	0.0002	0.0002	0.0005	0.0000	0.0001	0.0000	0.0001	0.0000	0.0002	0.0000	0.0000	0.0000	0.0004	0.0001	0.0004	0.0001	0.0006	0.0006	0.0000	0.0000	0.0000	0.0001	0.0002	0.0006	0.0000	0.0000	0.0000	0.0000	0.0000
я	0.0009	0.0003	0.0018	0.0002	0.0011	0.0003	0.0002	0.0002	0.0009	0.0001	0.0017	0.0008	0.0005	0.0017	0.0007	0.0000	0.0005	0.0015	0.0022	0.0002	0.0000	0.0000	0.0000	0.0004	0.0002	0.0002	0.0000	0.0000	0.0000	0.0000	0.0002
Hz	4.0613																														

Таблиці найбільш частих біграм для різних джерел:

Перехресні з пробілами:

о“ “ - 0.0245

е“ „ - 0.0178

и“ „ - 0.0174

„ „н - 0.0165

р„ „ - 0.0162

с“ „ - 0.0154

а“ „ - 0.0132

ь“ „ - 0.0106

„ „и - 0.0104

Не перехресні з пробілами:

о“ “ - 0.0240

и“ “ - 0.0179

е“ “ - 0.0172

“ “н - 0.0171

“ “п - 0.0154

“ “с - 0.0148

а“ “ - 0.0134

ь“ “ - 0.0111

“ “о - 0.0102

Перехресні без пробілів:

то — 0.0178

ак — 0.0166

но — 0.0141

на — 0.0112

ел — 0.0102

нж — 0.0102

Не перехресні без пробілів:

ак — 0.0163

ка — 0.0145

но — 0.0141

ов — 0.0133

не — 0.0129

ст — 0.0117

ко — 0.0111

ел — 0.0109

ни — 0.0100

Завдання 2

$H(10)$

Лабораторная работа №1

Произвольная часть текста:
ошибаться_в_определении_их_как_ошибаются_скажем_при_сложении_чисел_но_понят

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: 51

Поле ввода символов:

Продолжить Другой

Неравенство для энтропии:
 $1,69402212797767 < H < 2,38043048780011$

Двоичная таблица угаданных символов:
10000000000000000000000000000000
10000000000000000000000000000000
01000000000000000000000000000000
01000000000000000000000000000000
10000000000000000000000000000000

Вероятности:
q[1] = 0,5294117
q[2] = 0,1568627
q[3] = 0,0196078
q[4] = 0,1176470
q[5] = 0
q[6] = 0,0196078
q[7] = 0
q[8] = 0,0196078
q[9] = 0
q[10] = 0
q[11] = 0
q[12] = 0
q[13] = 0,0196078
q[14] = 0
q[15] = 0,0196078
q[16] = 0
q[17] = 0
q[18] = 0
q[19] = 0,0196078
q[20] = 0,0196078
q[21] = 0,0196078
q[22] = 0
q[23] = 0,0196078
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] = 0,0196078

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

H(20)

Лабораторная работа №1

Произвольная часть текста:
али_кусочек_апельсина_совсем_при_других_обстоятельствах_или_что_случилось_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: н

Символ по счету: 1

Номер эксперимента: 50

Поле ввода символов:
н

Продолжить Другой

Неравенство для энтропии:
 $1,50215130518827 < H < 2,27603171251419$

Двоичная таблица угаданных символов:
000000100000000000000000000000
000100000000000000000000000000
100000000000000000000000000000
001000000000000000000000000000
100000000000000000000000000000

Вероятности:
q[1] = 0,58
q[2] = 0,12
q[3] = 0,04
q[4] = 0,08
q[5] = 0
q[6] = 0,02
q[7] = 0,02
q[8] = 0,02
q[9] = 0,02
q[10] = 0,02
q[11] = 0
q[12] = 0
q[13] = 0
q[14] = 0
q[15] = 0
q[16] = 0
q[17] = 0,02
q[18] = 0
q[19] = 0,02
q[20] = 0
q[21] = 0
q[22] = 0
q[23] = 0,04
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

H(30)

Лабораторная работа №1

✕

Произвольная часть текста:

й_давай_ты_же_обещал_каждый_день_люди_произносят_подобное_как_образованные_

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: е

Символ по счету: 1

Номер эксперимента: 50

Поле ввода символов:

е

Продолжить

Другой

Неравенство для энтропии:

1,35537332981934 < H < 2,18006166784613

Двоичная таблица угаданных символов:

01000000000000000000000000000000

10000000000000000000000000000000

10000000000000000000000000000000

10000000000000000000000000000000

10000000000000000000000000000000

Вероятности:

q[1] = 0,62

q[2] = 0,1

q[3] = 0,08

q[4] = 0,02

q[5] = 0,02

q[6] = 0,02

q[7] = 0

q[8] = 0,02

q[9] = 0,02

q[10] = 0,02

q[11] = 0

q[12] = 0

q[13] = 0

q[14] = 0,02

q[15] = 0

q[16] = 0

q[17] = 0,02

q[18] = 0,02

q[19] = 0

q[20] = 0

q[21] = 0

q[22] = 0

q[23] = 0

q[24] = 0

q[25] = 0

q[26] = 0

q[27] = 0,02

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

Висновок:

Впродовж виконання лабораторної роботи ознайомились з поняттям ентропії на символ джерела та навчились її визначати, освоїли вимірювання частоти повторювання символів в тексті. Надлишковість може існувати в діапазоні 0.11-0.22. Також помітили за допомогою CoolPinkProgram що $H^{(n)}$ менше, чим більше порядок n-грамми.