

«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 буквених символів в текст (розмір від 1 Мб, визначення ентропії 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:
«Война и Мир» Льва Толстого

Джерело	H	R
H1 текст з пробілами	4.226	0.1548
H1 текст без пробілів	4.313	0.1374
H2 перехресна біграма з пробілами	3.7547	0.2491
H2 не перехресна біграма з пробілами	3.7532	0.2493
H2 перехресна біграма без пробілів	3.8824	0.2235
H2 не перехресна біграма без пробілів	3.8846	0.2230

Ч

Частоти літер у тексті з пробілами:

{ ' ': 0.17278987407351912,
'о': 0.0931007729298065,
'а': 0.06808690031910179,
'е': 0.06607299088407577,
'и': 0.05498125134123067,
'н': 0.05385874444301944,
'т': 0.04664244726820677,
'с': 0.043163183805425366,
'л': 0.041568309750953304,
'в': 0.037691597578737385,
'р': 0.03735509947011072,

'к': 0.029394697043768882,
'д': 0.024878003563070537,
'м': 0.024526267804241907,
'у': 0.023464711506838532,
'п': 0.021154937697057737,
'я': 0.018856592124928413,
'г': 0.01670554383808473,
'ь': 0.016088419004905253,
'з': 0.014371643748817494,
'б': 0.014303074322908664,
'ч': 0.011442205497490232,
'й': 0.009590830997951806,
'ж': 0.008351502114858867,
'ш': 0.007632792946999643,
'х': 0.006989002225966733,
'ю': 0.005316670116301365,
'ц': 0.0033319661774957685,
'щ': 0.002251362817339938,
'ф': 0.0017891540945470798}
H1 = 4.226

Частоти літер у тексті без пробілів:

{ 'о': 0.11265949811975479,
'а': 0.08238655136749996,
'е': 0.07995844558200005,
'и': 0.06653778240176544,
'н': 0.06517619903090917,
'т': 0.05643348142729591,
'с': 0.0522303995770793,
'л': 0.050303282706815444,
'в': 0.04560535904461721,
'р': 0.04520118700563843,
'к': 0.035567139430132795,
'д': 0.0301023646369444,
'м': 0.02968282483982954,
'у': 0.028387322828958388,
'п': 0.02559961150307432,
'я': 0.022816510504631075,
'г': 0.020217822603820733,
'ь': 0.019466339230966648,
'з': 0.017387081555155653,
'б': 0.017308705988661667,
'ч': 0.013844813304790284,
'й': 0.011607267720177714,

'ж': 0.010107374526096759,

'ㅍ': 0.00923448586396768,

'x': 0.00844765664740062,

'ю': 0.00643448033157475,

'Ц': 0.0040324997349061725

'щ': 0.0027247035175261597

'ϕ': 0.002165317121373017}

H1 = 4.313

Частоти перехресних біграм у тексті з пробілами

[illegible]

Частоти не перехресних біграм у тексті з пробілами

[illegible]

Частоти перехресних біграм у тексті без пробілів

	а	б	в	г	д	е	ж	з	и	к	л	м	н	о	п	р	с	т	у	ф	х	ц	ч	ш	щ	ъ	ы	э	ю	я
а	0.0003	0.0015	0.0056	0.0165	0.0032	0.0015	0.0016	0.0051	0.0015	0.0010	0.0065	0.0107	0.0040	0.0076	0.0014	0.0027	0.0035	0.0061	0.0063	0.0086	0.0006	0.0013	0.0002	0.0013	0.0012	0.0000	0.0000	0.0000	0.0000	0.0034
б	0.0013	0.0000	0.0001	0.0000	0.0000	0.0024	0.0000	0.0000	0.0009	0.0000	0.0003	0.0009	0.0000	0.0003	0.0027	0.0000	0.0014	0.0001	0.0000	0.0015	0.0000	0.0001	0.0000	0.0002	0.0001	0.0000	0.0000	0.0000	0.0006	
в	0.0067	0.0003	0.0004	0.0003	0.0007	0.0052	0.0000	0.0007	0.0042	0.0000	0.0007	0.0013	0.0003	0.0020	0.0007	0.0011	0.0008	0.0044	0.0000	0.0009	0.0000	0.0001	0.0001	0.0002	0.0015	0.0000	0.0002	0.0000	0.0002	
г	0.0012	0.0000	0.0002	0.0000	0.0011	0.0007	0.0000	0.0001	0.0010	0.0000	0.0002	0.0019	0.0000	0.0004	0.0105	0.0001	0.0015	0.0002	0.0000	0.0009	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	
д	0.0048	0.0001	0.0011	0.0001	0.0001	0.0051	0.0000	0.0001	0.0033	0.0000	0.0003	0.0008	0.0002	0.0020	0.0004	0.0002	0.0023	0.0005	0.0005	0.0019	0.0000	0.0001	0.0002	0.0001	0.0001	0.0000	0.0000	0.0006	0.0000	0.0004
е	0.0002	0.0022	0.0034	0.0046	0.0037	0.0020	0.0013	0.0020	0.0014	0.0026	0.0021	0.0074	0.0057	0.0096	0.0016	0.0031	0.0085	0.0062	0.0048	0.0006	0.0001	0.0014	0.0003	0.0016	0.0010	0.0007	0.0000	0.0000	0.0004	0.0005
ж	0.0017	0.0001	0.0000	0.0000	0.0008	0.0004	0.0000	0.0000	0.0013	0.0000	0.0001	0.0000	0.0000	0.0012	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
з	0.0062	0.0000	0.0000	0.0000	0.0009	0.0004	0.0002	0.0001	0.0004	0.0000	0.0003	0.0002	0.0003	0.0019	0.0000	0.0002	0.0002	0.0002	0.0001	0.0003	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0012	0.0001	0.0006
и	0.0004	0.0014	0.0052	0.0016	0.0030	0.0033	0.0006	0.0025	0.0021	0.0017	0.0034	0.0006	0.0040	0.0057	0.0021	0.0026	0.0014	0.0055	0.0045	0.0006	0.0001	0.0021	0.0017	0.0020	0.0006	0.0001	0.0000	0.0009	0.0003	0.0021
к	0.0002	0.0004	0.0007	0.0003	0.0007	0.0001	0.0001	0.0001	0.0002	0.0006	0.0000	0.0008	0.0004	0.0005	0.0010	0.0006	0.0009	0.0004	0.0018	0.0006	0.0002	0.0001	0.0001	0.0001	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001
л	0.0089	0.0007	0.0006	0.0002	0.0003	0.0006	0.0002	0.0001	0.0033	0.0003	0.0000	0.0005	0.0005	0.0002	0.0030	0.0102	0.0004	0.0022	0.0006	0.0007	0.0019	0.0000	0.0001	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001
м	0.0076	0.0004	0.0008	0.0004	0.0008	0.0048	0.0005	0.0002	0.0000	0.0000	0.0017	0.0002	0.0001	0.0011	0.0004	0.0009	0.0003	0.0024	0.0003	0.0015	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0042	0.0010	0.0021
н	0.0033	0.0003	0.0000	0.0004	0.0004	0.0004	0.0001	0.0001	0.0002	0.0047	0.0000	0.0007	0.0003	0.0003	0.0019	0.0047	0.0011	0.0003	0.0011	0.0003	0.0030	0.0001	0.0001	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0006
о	0.0119	0.0004	0.0000	0.0004	0.0015	0.0105	0.0001	0.0002	0.0005	0.0000	0.0005	0.0000	0.0001	0.0039	0.0103	0.0006	0.0002	0.0010	0.0010	0.0034	0.0000	0.0001	0.0007	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001
п	0.0002	0.0049	0.0110	0.0059	0.0061	0.0031	0.0023	0.0016	0.0022	0.0042	0.0032	0.0005	0.0056	0.0103	0.0002	0.0034	0.0078	0.0000	0.0009	0.0006	0.0007	0.0008	0.0002	0.0026	0.0015	0.0002	0.0000	0.0000	0.0000	0.0011
р	0.0014	0.0000	0.0000	0.0000	0.0000	0.0029	0.0000	0.0000	0.0000	0.0000	0.0001	0.0000	0.0000	0.0001	0.0100	0.0001	0.0072	0.0000	0.0001	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0004
с	0.0001	0.0002	0.0000	0.0000	0.0005	0.0071	0.0004	0.0001	0.0002	0.0000	0.0004	0.0001	0.0003	0.0010	0.0001	0.0002	0.0001	0.0005	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0004
т	0.0017	0.0002	0.0022	0.0001	0.0005	0.0054	0.0001	0.0001	0.0002	0.0000	0.0000	0.0007	0.0022	0.0013	0.0012	0.0003	0.0000	0.0004	0.0003	0.0010	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0013
у	0.0004	0.0003	0.0030	0.0002	0.0004	0.0054	0.0001	0.0001	0.0004	0.0000	0.0007	0.0000	0.0002	0.0015	0.0159	0.0005	0.0034	0.0015	0.0003	0.0022	0.0000	0.0001	0.0001	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0006
ф	0.0003	0.0009	0.0016	0.0016	0.0025	0.0003	0.0016	0.0011	0.0007	0.0001	0.0017	0.0019	0.0014	0.0011	0.0005	0.0013	0.0009	0.0022	0.0020	0.0001	0.0000	0.0000	0.0000	0.0012	0.0000	0.0000	0.0000	0.0000	0.0000	0.0002
х	0.0002	0.0000	0.0000	0.0000	0.0000	0.0002	0.0000	0.0000	0.0000	0.0008	0.0000	0.0000	0.0000	0.0001	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	0.0000	0.0000
ц	0.0011	0.0002	0.0004	0.0002	0.0002	0.0001	0.0000	0.0001	0.0005	0.0000	0.0002	0.0002	0.0002	0.0004	0.0029	0.0004	0.0002	0.0005	0.0001	0.0002	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
ч	0.0007	0.0000	0.0001	0.0000	0.0000	0.0013	0.0000	0.0000	0.0003	0.0000	0.0002	0.0000	0.0000	0.0000	0.0005	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
ш	0.0026	0.0000	0.0000	0.0000	0.0000	0.0003	0.0000	0.0000	0.0001	0.0015	0.0000	0.0003	0.0000	0.0000	0.0006	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
щ	0.0016	0.0000	0.0000	0.0000	0.0000	0.0026	0.0000	0.0000	0.0002	0.0000	0.0004	0.0005	0.0000	0.0004	0.0003	0.0001	0.0000	0.0000	0.0002	0.0003	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
ъ	0.0005	0.0000	0.0000	0.0000	0.0000	0.0001	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	0.0001	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
ы	0.0006	0.0006	0.0013	0.0003	0.0004	0.0017	0.0001	0.0004	0.0010	0.0000	0.0020	0.0002	0.0007	0.0023	0.0000	0.0010	0.0002	0.0021	0.0006	0.0002	0.0000	0.0001	0.0001	0.0000	0.0005	0.0000	0.0000	0.0000	0.0005	0.0005
э	0.0001	0.0005	0.0003	0.0001	0.0005	0.0001	0.0001	0.0001	0.0004	0.0000	0.0003	0.0001	0.0002	0.0004	0.0002	0.0003	0.0002	0.0004	0.0007	0.0001	0.0000	0.0000	0.0001	0.0002	0.0000	0.0001	0.0005	0.0000	0.0000	0.0000
ю	0.0003	0.0005	0.0016	0.0007	0.0013	0.0005	0.0006	0.0020	0.0010	0.0000	0.0011	0.0012	0.0008	0.0022	0.0000	0.0011	0.0003	0.0002	0.0022	0.0003	0.0001	0.0003	0.0001	0.0005	0.0000	0.0000	0.0000	0.0000	0.0000	0.0003

H2= 3.8824

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	ь	ю	я
а	0.0003	0.0015	0.0056	0.0015	0.0033	0.0015	0.0016	0.0052	0.0015	0.0010	0.0065	0.0107	0.0039	0.0077	0.0014	0.0026	0.0034	0.0062	0.0064	0.0005	0.0006	0.0013	0.0002	0.0014	0.0012	0.0003	0.0000	0.0009	0.0034
б	0.0014	0.0000	0.0001	0.0000	0.0000	0.0024	0.0000	0.0000	0.0009	0.0000	0.0003	0.0009	0.0001	0.0003	0.0027	0.0000	0.0015	0.0001	0.0000	0.0015	0.0000	0.0001	0.0000	0.0000	0.0000	0.0002	0.0001	0.0000	0.0006
в	0.0066	0.0003	0.0004	0.0003	0.0005	0.0050	0.0000	0.0000	0.0041	0.0000	0.0000	0.0014	0.0003	0.0020	0.0000	0.0011	0.0000	0.0044	0.0007	0.0010	0.0000	0.0001	0.0001	0.0002	0.0015	0.0000	0.0002	0.0000	0.0002
г	0.0012	0.0000	0.0002	0.0000	0.0011	0.0000	0.0000	0.0001	0.0010	0.0000	0.0002	0.0019	0.0000	0.0005	0.0106	0.0001	0.0016	0.0002	0.0000	0.0010	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
д	0.0047	0.0001	0.0012	0.0001	0.0051	0.0000	0.0000	0.0000	0.0033	0.0000	0.0004	0.0007	0.0002	0.0020	0.0043	0.0002	0.0023	0.0005	0.0004	0.0019	0.0000	0.0001	0.0002	0.0001	0.0001	0.0000	0.0006	0.0000	0.0004
е	0.0002	0.0021	0.0033	0.0045	0.0037	0.0020	0.0014	0.0020	0.0014	0.0027	0.0022	0.0075	0.0057	0.0096	0.0016	0.0031	0.0004	0.0062	0.0048	0.0006	0.0001	0.0013	0.0004	0.0015	0.0010	0.0007	0.0000	0.0004	0.0005
ж	0.0016	0.0001	0.0000	0.0000	0.0000	0.0043	0.0000	0.0000	0.0013	0.0000	0.0001	0.0000	0.0000	0.0012	0.0001	0.0000	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
з	0.0061	0.0002	0.0010	0.0006	0.0008	0.0004	0.0002	0.0001	0.0004	0.0000	0.0003	0.0003	0.0003	0.0019	0.0008	0.0002	0.0002	0.0003	0.0001	0.0003	0.0000	0.0000	0.0000	0.0000	0.0001	0.0000	0.0012	0.0001	0.0006
и	0.0004	0.0014	0.0054	0.0012	0.0029	0.0034	0.0007	0.0025	0.0021	0.0018	0.0034	0.0061	0.0041	0.0055	0.0022	0.0026	0.0013	0.0054	0.0045	0.0006	0.0001	0.0020	0.0016	0.0021	0.0006	0.0001	0.0000	0.0003	0.0021
й	0.0002	0.0004	0.0007	0.0003	0.0007	0.0001	0.0001	0.0002	0.0006	0.0000	0.0000	0.0003	0.0005	0.0010	0.0006	0.0008	0.0004	0.0018	0.0006	0.0002	0.0001	0.0001	0.0000	0.0004	0.0002	0.0000	0.0000	0.0000	0.0001
к	0.0008	0.0007	0.0006	0.0002	0.0003	0.0005	0.0002	0.0001	0.0032	0.0000	0.0004	0.0005	0.0002	0.0030	0.0101	0.0003	0.0023	0.0006	0.0007	0.0019	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001
л	0.0075	0.0004	0.0009	0.0004	0.0008	0.0046	0.0005	0.0002	0.0000	0.0000	0.0018	0.0002	0.0001	0.0011	0.0005	0.0009	0.0003	0.0023	0.0003	0.0015	0.0000	0.0000	0.0000	0.0007	0.0001	0.0000	0.0041	0.0010	0.0021
м	0.0033	0.0003	0.0000	0.0004	0.0003	0.0035	0.0002	0.0002	0.0048	0.0000	0.0007	0.0002	0.0003	0.0020	0.0049	0.0010	0.0003	0.0011	0.0003	0.0029	0.0001	0.0001	0.0000	0.0005	0.0001	0.0000	0.0000	0.0001	0.0005
н	0.0120	0.0004	0.0005	0.0004	0.0014	0.0106	0.0001	0.0002	0.0005	0.0000	0.0005	0.0001	0.0001	0.0040	0.0102	0.0006	0.0002	0.0011	0.0009	0.0033	0.0000	0.0001	0.0007	0.0004	0.0001	0.0002	0.0011	0.0002	0.0037
о	0.0002	0.0049	0.0119	0.0059	0.0060	0.0030	0.0022	0.0016	0.0023	0.0043	0.0032	0.0004	0.0065	0.0103	0.0023	0.0035	0.0000	0.0097	0.0009	0.0006	0.0007	0.0000	0.0002	0.0026	0.0015	0.0002	0.0000	0.0000	0.0011
п	0.0014	0.0000	0.0000	0.0000	0.0000	0.0031	0.0000	0.0000	0.0000	0.0000	0.0001	0.0000	0.0000	0.0001	0.0100	0.0001	0.0072	0.0000	0.0001	0.0007	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0006	0.0000	0.0004
р	0.0005	0.0002	0.0004	0.0003	0.0005	0.0072	0.0004	0.0001	0.0062	0.0000	0.0004	0.0001	0.0003	0.0010	0.0000	0.0003	0.0001	0.0004	0.0009	0.0035	0.0000	0.0001	0.0001	0.0001	0.0003	0.0001	0.0010	0.0002	0.0012
с	0.0017	0.0002	0.0023	0.0001	0.0005	0.0043	0.0001	0.0001	0.0020	0.0000	0.0056	0.0028	0.0013	0.0012	0.0033	0.0021	0.0004	0.0012	0.0131	0.0011	0.0001	0.0002	0.0000	0.0004	0.0001	0.0000	0.0038	0.0001	0.0037
т	0.0064	0.0003	0.0030	0.0002	0.0004	0.0054	0.0001	0.0001	0.0040	0.0000	0.0007	0.0004	0.0003	0.0015	0.0159	0.0005	0.0035	0.0015	0.0004	0.0022	0.0000	0.0001	0.0001	0.0004	0.0000	0.0000	0.0059	0.0001	0.0006
у	0.0002	0.0000	0.0016	0.0017	0.0025	0.0003	0.0015	0.0011	0.0007	0.0001	0.0017	0.0019	0.0013	0.0011	0.0005	0.0013	0.0010	0.0022	0.0020	0.0001	0.0000	0.0004	0.0000	0.0012	0.0009	0.0002	0.0000	0.0012	0.0002
ф	0.0002	0.0000	0.0000	0.0000	0.0000	0.0002	0.0000	0.0000	0.0000	0.0007	0.0000	0.0000	0.0001	0.0000	0.0000	0.0006	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
х	0.0012	0.0001	0.0005	0.0002	0.0001	0.0001	0.0001	0.0001	0.0005	0.0000	0.0002	0.0003	0.0002	0.0004	0.0028	0.0004	0.0003	0.0005	0.0002	0.0002	0.0000	0.0000	0.0000	0.0000	0.0001	0.0000	0.0000	0.0000	0.0000
ц	0.0007	0.0000	0.0000	0.0000	0.0000	0.0013	0.0000	0.0000	0.0003	0.0000	0.0002	0.0000	0.0000	0.0000	0.0005	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
ч	0.0026	0.0000	0.0000	0.0000	0.0000	0.0035	0.0000	0.0000	0.0015	0.0000	0.0003	0.0001	0.0000	0.0006	0.0001	0.0000	0.0000	0.0000	0.0038	0.0007	0.0000	0.0000	0.0000	0.0000	0.0001	0.0000	0.0002	0.0000	0.0000
ш	0.0016	0.0000	0.0000	0.0000	0.0000	0.0026	0.0000	0.0000	0.0022	0.0000	0.0004	0.0005	0.0000	0.0005	0.0003	0.0001	0.0000	0.0000	0.0002	0.0003	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0004	0.0000	0.0000
щ	0.0005	0.0000	0.0000	0.0000	0.0000	0.0014	0.0000	0.0000	0.0008	0.0000	0.0000	0.0000	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	0.0000	0.0000
ь	0.0006	0.0006	0.0013	0.0003	0.0004	0.0017	0.0001	0.0005	0.0009	0.0000	0.0021	0.0002	0.0007	0.0023	0.0008	0.0010	0.0002	0.0022	0.0006	0.0002	0.0001	0.0001	0.0001	0.0005	0.0005	0.0000	0.0000	0.0005	0.0006
ю	0.0001	0.0005	0.0003	0.0001	0.0006	0.0001	0.0001	0.0001	0.0004	0.0000	0.0003	0.0001	0.0002	0.0003	0.0002	0.0003	0.0002	0.0005	0.0007	0.0001	0.0000	0.0000	0.0001	0.0002	0.0001	0.0005	0.0000	0.0000	0.0000
я	0.0003	0.0005	0.0015	0.0007	0.0013	0.0004	0.0007	0.0020	0.0010	0.0000	0.0011	0.0013	0.0000	0.0022	0.0000	0.0011	0.0003	0.0020	0.0023	0.0003	0.0000	0.0002	0.0001	0.0005	0.0001	0.0002	0.0000	0.0001	0.0003

H2= 3.8846

Найбільш часті біграми для різних джерел:

Перехресні з пробілами:

о «» - 0.0200

и « » - 0.0173

е «» - 0.0156

а «» - 0.0152

«» «» - 0.0151

«» с - 0.0151

«» п - 0.0149

«» н - 0.0144

«» о - 0.0115

«» и - 0.0106

Не перехресні з пробілами:

«» о – 0.0200

«» и – 0.0171

«» е – 0.0154

«» а – 0.0154

р «» – 0.0152

п «» – 0.0151

«» «» – 0.0150

в «» – 0.0148

н «» – 0.0143

Перехресні без пробілів:

от – 0.0159

тс – 0.0129

ан – 0.0119

во – 0.0118

ла – 0.0107

ен – 0.0105

ог – 0.0105
он – 0.0103
ок – 0.0102

Не перехресні без пробілів:

от – 0.0159
тс – 0.0131
ан – 0.0120
во – 0.0119
ла – 0.0107
ог – 0.0106
ен – 0.0106
но – 0.0103
он – 0.0102

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] = 0,019607
q[14] = 0
q[15] = 0,019607
q[16] = 0
q[17] = 0
q[18] = 0
q[19] = 0,019607
q[20] = 0,019607
q[21] = 0,019607
q[22] = 0
q[23] = 0,019607
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] = 0,019607

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

	R(з пробілами)	R(без пробілів)
H1	0.1254	0.1135
H2	0.2026	0.1797
H2 (крок 2)	0.2032	0.1798

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.