

Міністерство освіти і науки України
Національний технічний університет України
"Київський політехнічний інститут імені Ігоря Сікорського"
Фізико-технічний інститут

КРИПТОГРАФІЯ
КОМП'ЮТЕРНИЙ ПРАКТИКУМ №1

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

Виконали:
Студенти ФБ-01
Сотнікова П.О.
Струкало В.В.

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.

Частоти букв у тексті з пробілами

Letters Frequency	
' '	0.16763719112538508
'а'	0.06363574916431802
'б'	0.015144687684341613
'в'	0.036055171396735926
'г'	0.014219907255685916
'д'	0.026395670839614603
'е'	0.07060488628170676
'ж'	0.00903478731074261
'з'	0.013541939765353608
'и'	0.05865852559480894
'й'	0.00826055253326342
'к'	0.02672338926394442
'л'	0.03767123287671233
'м'	0.026970202202267812
'н'	0.05374070098970964
'о'	0.09874156125057351
'п'	0.02093813495444714
'р'	0.03296642524742741
'с'	0.04704295569246903
'т'	0.05520314445828144
'у'	0.023155354919053548
'ф'	0.0005345906796880121
'х'	0.00821446712984204
'ц'	0.0031348315527298947
'ч'	0.01374369142033165
'ш'	0.0060720079307858684
'щ'	0.0029801894212492627
'ъ'	0.0002888018614406502
'ы'	0.015204086648751393
'ь'	0.01749299501867995
'э'	0.0027098217211771647
'ю'	0.005929655240217605
'я'	0.017345521727731534
'ё'	7.168840532214721e-06

Частоти букв у тексті без пробілів

Frequency of Letters	
'а'	0.07645193716710612
'б'	0.01819481543729445
'в'	0.04331665352213879
'г'	0.017083784984232717
'д'	0.03171173742769997
'е'	0.08482465281836313
'ж'	0.010854386109535544
'з'	0.016269275394059
'и'	0.07047230482836855
'й'	0.009924221079065258
'к'	0.03210545807551808
'л'	0.04525818846669185
'м'	0.032401978938406094
'н'	0.06456403435704804
'о'	0.11862803118759682
'п'	0.025155058264503993
'р'	0.03960583641645309
'с'	0.05651736861726539
'т'	0.06632101274793635
'у'	0.027818824522398398
'ф'	0.0006422568067532934
'х'	0.009868854112965836
'ц'	0.003766184071785117
'ч'	0.016511659667872026
'ш'	0.007294905377854936
'щ'	0.003580397141095946
'ъ'	0.00034696632088971024
'ы'	0.018266177304711483
'ь'	0.0210160699543161
'э'	0.0032555776066460046
'ю'	0.007123882971458944
'я'	0.02083889566279795
'ё'	8.612639171021176e-06

Таблиці частот біграм наведено в таблиці

Ентропія

Entropy	4.3547600517319545				
H1 w/ spaces	4.3547600517319545	surplus:	0.14402125624927598		
H1 w/o spaces	4.448167298042461	surplus:	0.11819592347630059		
H2 w/ spaces	3.9436132206445955	surplus:	0.2248369484551599		
H2 w/o spaces	4.126240277770246	surplus:	0.18201469192597075		
H2 w/ spaces crossing	3.9438360468685425	surplus:	0.2247931493688765		
H2 w/o spaces crossing	4.127100273584213	surplus:	0.18184420647348276		

«найчастіші» букви, біграми

Текст з пробілами

' '	0.16763719112538508
'о'	0.09874156125057351
'е'	0.07060488628170676
'а'	0.06363574916431802
'и'	0.05865852559480894
'т'	0.05520314445828144
'н'	0.05374070098970964
'с'	0.04704295569246903
'л'	0.03767123287671233
'в'	0.036055171396735926

З перехрестям

' а'	0.0023636691354787967
'ау'	9.933964737497542e-05
'лч'	9.831552729894475e-05
'пъ'	9.831552729894475e-05
'ье'	9.524316707085273e-05
'бщ'	9.421904699482204e-05
'тд'	9.421904699482204e-05
'жж'	9.21708068427607e-06
'йи'	9.21708068427607e-06
'пс'	9.21708068427607e-06

Без перехрестя

'а'	9.831552729894475e-05
'б'	9.831552729894475e-05
'в'	9.831552729894475e-05
'г'	9.62672871468834e-05
'д'	9.62672871468834e-05
'е'	9.62672871468834e-05
'ж'	9.62672871468834e-05
'з'	9.421904699482204e-05
'и'	9.012256669069935e-05
'й'	9.012256669069935e-05

Текст без пробілів

'я'	0.02083889566279795
'ю'	0.007123882971458944
'э'	0.0032555776066460046
'ь'	0.0210160699543161
'ы'	0.018266177304711483
'ъ'	0.00034696632088971024
'щ'	0.003580397141095946
'ш'	0.007294905377854936
'ч'	0.016511659667872026
'ц'	0.003766184071785117

З перехрестя

'аа'	0.0003248195344499415
'аб'	0.0013165034161418084
'ав'	0.005091300127097947
'аг'	0.0014862954455133687
'ад'	0.0029221454330250417
'ае'	0.0021937622345615366
'аж'	0.0014506145118048524
'аз'	0.004408440878538411
'аи'	0.0017397531125462776
'ай'	0.0009277042764214238

Без перехрестя

'аа'	0.0003125153796938826
'гш'	9.843004084846695e-06
'жж'	9.843004084846695e-06
'жз'	9.843004084846695e-06
'жр'	9.843004084846695e-06
'нш'	9.843004084846695e-06
'тф'	9.843004084846695e-06
'хф'	9.843004084846695e-06
'цт'	9.843004084846695e-06
'эн'	9.843004084846695e-06

CoolPinkProgram

$$H(10) = 4,7399$$

Лабораторная работа №1

[illegible]

$$H(20) = 2,0717$$

Лабораторная работа №1

×

Произвольная часть текста:	
ть_на_землю_или_не_упасть_будучи_организмом_человек_должен_подчиняться_разл	
Использованные буквы:	
е, ы, _, с, л, д, в, а, ж, э, ш, я, ю, й, ф, х,	
Порядок n-граммы:	Введенный символ: у
5 символов 10 символов 15 символов 20 символов 25 символов 30 символов 35 символов 40 символов 45 символов 50 символов	Символ по счету: 18
	Номер эксперимента: 51
	Поле ввода символов: у
	<button>Продолжить</button> <button>Другой</button>
Неравенство для энтропии: 1,77175623391097 < H < 2,37173959443901	
Двоичная таблица угаданных символов: 10000000000000000000000000000000 ^ 00010000000000000000000000000000 00000000000000000000000000000000 1000 000000000000000000000000010000000000 01000000000000000000000000000000 v ~~~~~	
Строка состояния: Вы угадали. Для продолжения опыта нажмите "Продолжить", или "Другой" для выбора другого порядка	

$$H(30) = 2,6446$$

Лабораторная работа №1

✕

[illegible]

Висновки: на лабораторній роботі познайомилися з поняттями ентропії та надлишковості, вивчали та порівнювали різні моделі джерела відкритого тексту для наближеного визначення ентропії. Також визначали частоту символів та біграм в тексті та оцінювали ентропію.