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

Частоти букв у тексті без пробілів

Letters Frequency		
- 11	0.16763719112538508	
'a'	0.06363574916431802	
'6'	0.015144687684341613	
'в'	0.036055171396735926	
161	0.014219907255685916	
'A'	0.026395670839614603	
'e'	0.07060488628170676	
'ж'	0.00903478731074261	
'3'	0.013541939765353608	
'и'	0.05865852559480894	
'й'	0.00826055253326342	
'ĸ'	0.02672338926394442	
'л'	0.03767123287671233	
'm'	0.026970202202267812	
'H'	0.05374070098970964	
'0'	0.09874156125057351	
'n'	0.02093813495444714	
'p'	0.03296642524742741	
'c'	0.04704295569246903	
Ψ'	0.05520314445828144	
'y'	0.023155354919053548	
'φ'	0.0005345906796880121	
'x'	0.00821446712984204	
'ц'	0.0031348315527298947	
'4'	0.01374369142033165	
'ш'	0.0060720079307858684	
'щ'	0.0029801894212492627	
'ъ'	0.0002888018614406502	
'ы'	0.015204086648751393	
'ь'	0.01749299501867995	
'ə'	0.0027098217211771647	
'ю'	0.005929655240217605	
'я'	0.017345521727731534	
'ë'	7.168840532214721e-06	

Frequency of Letters			
'a'	0.07645193716710612		
'6'	0.01819481543729445		
'в'	0.04331665352213879		
'r'	0.017083784984232717		
'д'	0.03171173742769997		
'e'	0.08482465281836313		
'ж'	0.010854386109535544		
'3'	0.016269275394059		
'и'	0.07047230482836855		
'й'	0.009924221079065258		
'k'	0.03210545807551808		
'л'	0.04525818846669185		
'M'	0.032401978938406094		
'н'	0.06456403435704804		
'o'	0.11862803118759682		
'n'	0.025155058264503993		
'p'	0.03960583641645309		
'c'	0.05651736861726539		
' †'	0.06632101274793635		
'y'	0.027818824522398398		
'φ'	0.0006422568067532934		
'x'	0.009868854112965836		
'ц'	0.003766184071785117		
'4'	0.016511659667872026		
'ш'	0.007294905377854936		
'щ'	0.003580397141095946		
'ъ'	0.00034696632088971024		
'ы'	0.018266177304711483		
'ь'	0.0210160699543161		
'a'	0.0032555776066460046		
'ю'	0.007123882971458944		
'я'	0.02083889566279795		
'ë'	8.612639171021176e-06		

Таблиці частот біграм наведено в таблиці

Ентропія

Entropy	4.3547600517319545				
H1 w/ spaces	4.3547600517319545	surplus:	0.144023	125624927	598
H1 w/o spaces	4.448167298042461	surplus:	0.11819	592347630	059
H2 w/ spaces	3.9436132206445955	surplus:	0.224836	594845515	99
H2 w/o spaces	4.126240277770246	surplus:	0.182014	469192597	075
H2 w/ spaces crossing	3.9438360468685425	surplus:	0.224793	314936887	65
H2 w/o spaces crossing	4.127100273584213	surplus:	0.181844	120647348	276

«найчастіші» букви, біграми

Текст з пробілами

11	0.16763719112538508
'o'	0.09874156125057351
'e'	0.07060488628170676
'a'	0.06363574916431802
'и'	0.05865852559480894
'τ'	0.05520314445828144
'н'	0.05374070098970964
'c'	0.04704295569246903
'л'	0.03767123287671233
'в'	0.036055171396735926

3 перехрестям

'ay'	9.933964737497542e-05
'лч'	9.831552729894475e-05
'пь'	9.831552729894475e-05
'ъе'	9.524316707085273e-05
'бщ'	9.421904699482204e-05
'тд'	9.421904699482204e-05
'жж'	9.21708068427607e-06
'йи'	9.21708068427607e-06

'пс'	9.21708068427607e-06
'пч'	9.21708068427607e-06

Без перехрестя

'тд'	9.831552729894475e-05
'ыд'	9.831552729894475e-05
'яч'	9.831552729894475e-05
'щ'	9.62672871468834e-05
'кж'	9.62672871468834e-05
'шо'	9.62672871468834e-05
'ъе'	9.62672871468834e-05
'ay'	9.421904699482204e-05
'вк'	9.012256669069935e-05
'3ь'	9.012256669069935e-05

Текст без пробілів

'o'	0.11862803118759682
'e'	0.08482465281836313
'a'	0.07645193716710612
'и'	0.07047230482836855
'т'	0.06632101274793635
'н'	0.06456403435704804
'c'	0.05651736861726539
'л'	0.04525818846669185
'в'	0.04331665352213879
'p'	0.03960583641645309

3 перехрестя

'дм'	9.966053897895933e-05
'сж'	9.966053897895933e-05

'хч'	9.966053897895933e-05
'щу'	9.966053897895933e-05
'рз'	9.843016195452773e-05
'хг'	9.843016195452773e-05
'бп'	9.843016195452772e-06
'гэ'	9.843016195452772e-06
'кф'	9.843016195452772e-06
'хф'	9.843016195452772e-06

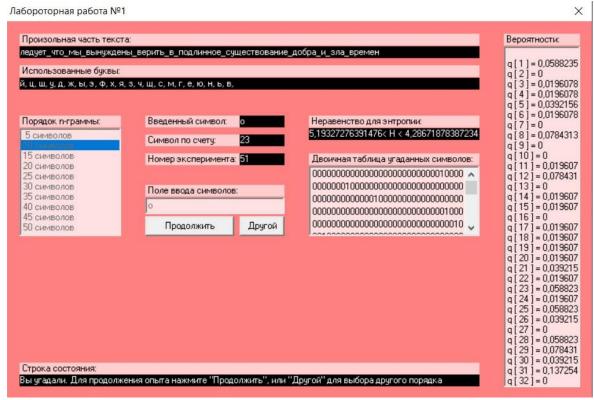
Без перехрестя

'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
'кг'	9.843004084846695e-05

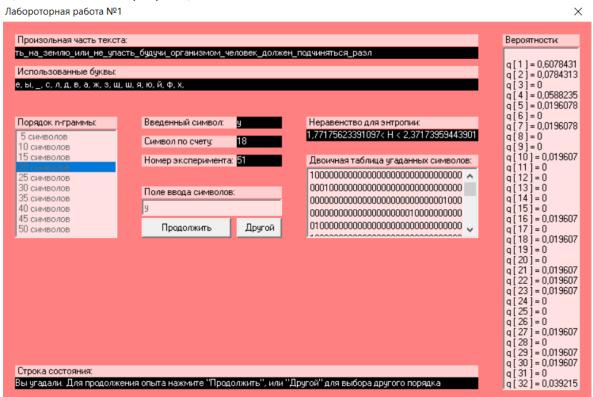
CoolPinkProgram

4,28671878387234< H(10) < 5,19327276391476

-0.3865455278< R(10) < 0.14265624322

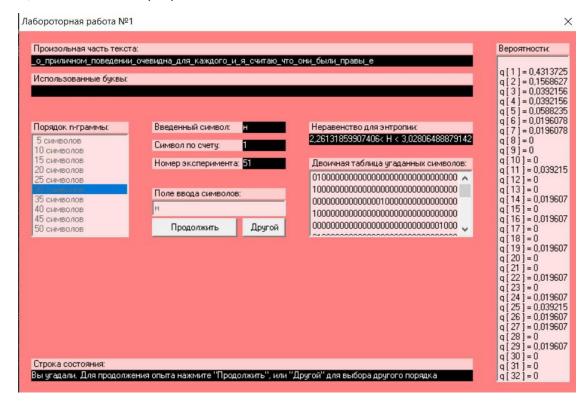


- 1,77175623391097 < H(20) < 2,37173959443901
- 0,525652082 < R(20) < 0,645648754



2,26131859907406 < H(30) < 3,02806488879142

0,394387022 < R(30) < 0.54773628018



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.